PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ :			(11) International Publication Number:	WO 94/12641	
C12N 15/31, C07K 13/00, A61K 39/ // (C12N 15/31, C12R 1:21)	/102	A1	(43) International Publication Date:	9 June 1994 (09.06.94)	
(21) International Application Number: PCT/CA93/00501 (22) International Filing Date: 23 November 1993 (23.11.93)			(81) Designated States: AU, BR, CA, FI, JP, KR, NO, NZ, RU, UA, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).		
(30) Priority Data: 9224584.4 23 November 1992 ((23.11.92	2) G	Published With international search report.		
(71) Applicant (for all designated States except US): LABORATORIES LIMITED [CA/CA]; 1 enue West, Willowdale, Ontario M2R 3T4	755 Stee	AUGH eles Av			
(72) Inventors; and (75) Inventors/Applicants (for US only): CHONG 32 Estoril Street, Richmond Hill, Ontario THOMAS, Wayne [AU/AU]; 31 Taylor I W.A. 6009 (AU). YANG, Yan-Ping [CN/0 120 Torresdale Avenue, Willowdale, On (CA). LOOSMORE, Sheena [CA/CA]; 70 Drive, Aurora, Ontario L4G 4R4 (CA). S. Charles [CA/CA]; 189 Mabley Crescent, TI L4J 2Z7 (CA). KLEIN, Michel [CA/CA]; 1 vard, Willowdale, Ontario M2P 1B9 (CA).	LAC 0B Road, No CA]; Api tario M2 Crawfor IA, Dwo hornhill, 16 Munro	6 (CA) edlands t. 1709 2R 3N rd Ros o, Yuan Ontario Boule	·		
(74) Agent: STEWART, Michael, I.; Sim & University Avenue, Suite 701, Toronto, Or (CA).	McBurne ntario M	ey, 336 5G 1R			
(54) Title: HAEMOPHILUS OUTER MEMBRAN	E PROT	EIN			
(57) Abstract			DIS SEQUENCE COMPRISION		
Purified and isolated nucleic acid from specific strains of Haemophilus influenzae is provided which encodes at least a portion of the D15 outer membrane protein of Haemophilus. The nucleic acid is used to produce peptides, polypeptides and proteins free of contaminant associated with Haemophilus for purposes of diagnosis and medical treatment. Furthermore, the nucleic acid may be used in the diagnosis of Haemophilus infection. Antisera obtained following immunization with the nucleic acid D15 outer membrane protein or peptides also may be used for the purpose of diagnosis and medical treatment.	SVIPTEM I	UKONLDANSZ NO EGYGFEI GLGENGNGSI	ALPHVARDIRADAQCILEQQIRASLARACQRVIINDVANIVRSLEVSGEFIDAGUQEDIA ALPHVARDIRETARSARBAYASVERARIVERTARILERARIASLIQUIRETERARIASLITY LOSDROVIINOMAAQCIDSDUQLAREETRARVIIDARELQADLASARI DAVGOSSARII A. A. A. A. B. B. B. H.	Zegen Minna Sell BMK CRESVSSSTLGEDRELGED Ca Engen Minna Sell BMK CA ENGEN	
	MOSAS MOSAS	VOJEVELER DESONIVE	PRINCEDECICIONESCISA(ASACOPELITICAAVSIACITODIXIDSALLOYEPYFINICI T.I. ADRISTAVLICATINICISEPALEORRE/TIQSOFRE/LUCINO ESPERANCSILERCITY EMANSIOSICIONESCISTE PRINTI DICUICE PERANCSILERA I PARACOCESTICITY R	Eagan MinnA SSIJ FAK DED/RASLOCKVTTPOSENCY Ca Regen MinnA SSIJ FAK CESSOVIERPURENSALIV Ca Eagan MinnA SSIJ	

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
ΑU	Australia	GE	Georgia	MW	
BB	Barbados	GN	Guinea		Malawi
BE	Belgium	GR	Greece	NE	Niger
BF	Burkina Faso	au		NL	Netherlands
BG	Bulgaria	_	Hungary	NO	Norway
BJ	-	Œ	Ireland	NZ	New Zealand
_	Benin	П	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic	SD	Sudan
CG	Congo		of Korea	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhatan	SK	
CM	Cameroon	u	Liechtenstein		Slova <u>kia</u>
CN	China	LK	Sri Lanka	SN	Senegal
CS	Czechoslovakia	LU		TD	Chad
CZ	Czech Republic	_	Luxembourg	TG	Togo
DE	-	LV		TJ	Tajikistan
	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
Fī	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Vict Nam
GA	Gabon		· - v -	* 1.5	A SING TARRET

1

TITLE OF INVENTION HAEMOPHILUS OUTER MEMBRANE PROTEIN

FIELD OF INVENTION

The present invention is related to the field of molecular genetics and is particularly concerned with the cloning of an outer membrane protein D15 of Haemophilus.

BACKGROUND OF THE INVENTION

Haemophilus influenzae type b (Hib) is a major cause of bacterial meningitis in children under the age of five 10 years. Protective antibodies to the disease are induced by the capsular polysaccharide of the organism and a vaccine was developed that utilises the purified polyribosyl ribitol phosphate (PRP) as the antigen. This vaccine provides 90% protection in adults and in children 15 over 24 months of age, but was ineffective in children under 24 months Zangwill et al 1993 (The references are identified in a list of reference at the end of this disclosure). Like other polysaccharide antigens, PRP 20 does not induce the proliferation of T-helper cells, and re-immunisation fails to elicit either a booster response or an increase in memory cells. Conjugation of the PRP polysaccharide with protein carriers confers T-cell dependent characteristics to the vaccine and 25 substantially enhances the immunologic response to the PRP antigen. Currently, there are four PRP-carrier conjugate vaccines available. These are vaccines based upon <u>H. influenzae</u> type b capsular polysaccharide conjugated to diphtheria toxoid, tetanus toxoid, 30 Neisseria meningitidis outer membrane protein (reviewed in Zangwill et al 1993).

However, the current <u>Haemophilus</u> conjugate vaccines only protect against meningitis caused by <u>Haemophilus</u> influenzae type b. They do not protect against other invasive typeable strains (types a and c) and, more importantly, against non-typeable (NTHi) strains which are a common cause of postpartum and neonatal sepsis,

pneumonia and otitis media. In the United States alone, treatment of otitis media costs between 1 and 2 billion dollars per year for antibiotics and surgical procedures, such as tonsillectomies, adenoidectomies and insertion of tympanostomy tubes. To achieve universal protection against H. influenzae related diseases in the 2 to 6 month age group and certain high risk groups, provision of conserved, cross-reactive non-capsular H. influenzae immunogens is desirable. Methods 10 immunity inducing against disease are constantly improving and there is presently a move to use subunits and better defined materials as antigens. This is being undertaken to minimise or eliminate potential sideeffects caused by certain native immunogens, preserving their immunogenicity to confer protection 15 against the disease. Therefore, it would be very attractive to develop a universal vaccine against <u>Haemophilus</u> using cross-reactive outer membrane proteins, fragment, analogs, and/or peptides corresponding thereto 20 protective antigens. Such antigens incorporated into the conventional H. influenzae type b conjugate vaccines as additional immunogens or used as autologous carriers for H. influenzae capsular polysaccharides. A high molecular weight outer membrane 25 protein D15 found in non-typeable and type b stains of H. influenzae has been identified as a cross-reactive antigen (Thomas et al., 1990). D15 appears to be cell surface-exposed in its natural state and exhibits a molecular mass of about 80 kDa as judged by SDS-PAGE analysis. It would be desirable to provide the sequence 30 of the DNA molecule that encodes this D15 outer membrane protein and peptides corresponding to portions thereof for diagnosis, immunization and the generation diagnostic and immunological reagents. The diseases caused by <u>Haemophilus</u> are serious and improved methods 35 for preventing, detecting and treating diseases such as

10

15

20

25

3

otitis media, epiglottitis, pneumonia, and tracheobronchitis, are required.

SUMMARY OF THE INVENTION

The present invention is directed towards the provision of purified and isolated nucleic acid molecules comprising at least a portion coding for a D15 outer membrane protein of a species of Haemophilus. The nucleic acid molecules comprising at least a portion coding for D15 outer membrane protein are useful for the specific detection of strains of Haemophilus, and for diagnosis of infection by Haemophilus. The purified and isolated nucleic acid molecules, such as DNA comprising at least a portion coding for D15 outer membrane protein, are also useful for expression of the D15 gene by recombinant DNA means for providing, in an economical manner, purified and isolated D15 outer membrane protein.

The D15 outer membrane protein or fragments thereof or analogs thereof are useful immunogenic compositions for the preparation of vaccines against diseases caused by <u>Haemophilus</u>, the diagnosis of infection by <u>Haemophilus</u> and as tools for the generation of immunological reagents. Mono- or polyclonal antisera (antibodies) raised against the D15 outer membrane protein produced in accordance with aspects of the present invention are useful for the diagnosis of infection by <u>Haemophilus</u>, specific detection of <u>Haemophilus</u> (in, for example, <u>in vitro</u> and <u>in vivo</u> assays) and for the treatment of diseases caused by infection by <u>Haemophilus</u>.

Peptides corresponding to portions of the D15 outer

membrane protein or analogs thereof are useful immunogenic compositions for the preparation of vaccines against disease caused by Haemophilus, the diagnosis of infection by Haemophilus and as tools for the generation of immunological reagents. Mono- or polyclonal antisera raised against these peptides, produced in accordance with aspects of the present invention, are useful for the

10

15

20

25

35

diagnosis of infection by Haemophilus, specific detection of <u>Haemophilus</u> (in, for example, <u>in vitro</u> and <u>in vivo</u> and for use in passive immunization as a assays) treatment of disease caused by infection by Haemophilus.

In accordance with one aspect of the present invention, therefore, there is provided a purified and isolated nucleic acid molecule, the molecule comprising at least a portion coding for a D15 outer membrane The nucleic acid molecule has a DNA sequence protein. selected from:

- (a) the DNA sequence set out in any one of Figures 1A to 1E (as described below) or its complementary strand; and
- (b) DNA sequences which hybridize under stringent conditions to the DNA sequences defined in (a). The DNA sequences defined in (b) preferably has at least 90% sequence identity with the sequences defined in (a). DNA sequence defined in (b) particularly may comprise the consensus sequence set forth in Figure 1F (as described below).

In another aspect of the present invention, there is provided a purified and isolated D15 outer membrane protein or a portion thereof. The D15 outer membrane protein may be a <u>Haemophilus</u> D15 outer membrane protein and more particularly an H. influenzae D15 outer membrane protein and the H. influenzae strain may be an H. influenzae type b strain, such as H. influenzae type b strains Ca or Eagan or MinnA or a non-typeable H. influenzae strain, such as PAK 12085 or SB33.

30 In an additional embodiment, the present invention includes a recombinant plasmid adapted transformation of a host, the recombinant plasmid comprising a plasmid vector into which has been inserted a DNA segment comprising the purified and isolated DNA molecule provided herein. Such recombinant plasmid comprises a plasmid vector into which a DNA segment which

10

15

20

25

30

35

comprises at least an 18 bp fragment selected from the DNA molecules as recited above is inserted. The recombinant plasmid may be plasmid DS-712-2-1 having ATCC accession number 75604, deposited November 4, 1993 and plasmid JB-1042-5-1 having ATCC accession number 75006, deposited November 4, 1993.

The plasmids may be adapted for expression of the encoded D15 outer membrane protein in a host cell, which may be a heterologous or homologous host, incorporation into a recombinant vector, provided in accordance with a further aspect of the invention. recombinant vector may comprise at least a DNA segment comprising at least an 18 bp fragment selected from the DNA molecules as recited above and expression means operatively coupled to the DNA segment for expression of the gene product encoded thereby in the host cell. plasmid for expression of the encoded D15 outer membrane protein may be plasmid DS-880-1-2 having ATCC accession number 75605, deposited November 4, 1993 being adapted for expression at the D15 outer membrane protein in E. The selected DNA segment may encode a polypeptide of at least 6 residues and, in particular, may be selected from those segments encoding a polypeptide of Table 2 (below). The DNA segment may further comprise a nucleic acid sequence encoding a leader sequence for export of the gene product from the host. The host for expression may be selected from, for example, Escherichia coli, Bacillus, Haemophilus, fungi, yeast the baculovirus expression system may be used.

Additional aspects of the invention include the protein encoded by the DNA molecule comprising at least a portion coding for the D15 outer membrane protein, fragment or a functional analog of such protein, the use of the protein or analog in vaccination and diagnosis, and the generation of immunological reagents. The invention also includes antisera (antibodies) raised

against the D15 outer membrane protein encoded by the DNA molecule comprising at least a portion coding for a D15 outer membrane protein and purified peptides corresponding to portions of the D15 outer membrane protein and there are in passive immunization and treatment of diseases caused by <u>Haemophilus</u>.

According to another aspect of the invention, a purified and isolated peptide containing an amino acid sequence corresponding to the amino acid sequence of at least a portion of the D15 outer membrane protein or variant or mutant which retains immunogenicity. The peptide may be produced by recombinant methods or peptide synthesis whereby the purified peptide is free from contaminants associated with bacteria normally containing the D15 outer membrane protein. Such synthetic peptides preferably have an amino acid sequence selected from those presented in Table 2.

In accordance with an additional aspect of the invention, an immunogenic composition is provided which comprises the D15 outer membrane protein, fragments thereof, functional analogs thereof, or peptides as recited above and a physiologically-acceptable carrier Such immunogenic composition is particularly therefor. formulated as a vaccine for in vivo administration to protect against diseases caused by Haemophilus. For such purpose, the immunogenic composition may be formulated as a microparticle preparation, capsule preparation or liposome preparation. In addition, such immunogenic composition may be provided in combination with a targeting molecule for delivery to specific cells of the immune system or to mucosal surfaces.

In accordance with a further aspect of the invention, there is provided a method for inducing protection against disease caused by <u>Haemophilus</u>, comprising the step of administering to a subject, including a mammal, such as a human, an effective amount

5

10

15

20

25

30

10

20

25

30

of the immunogenic composition or the nucleic acid molecule as recited above to provide protective immunity against <u>Haemophilus</u> infection.

The present invention further includes a chimeric comprising molecule a D15 protein orpeptide corresponding thereto as provided herein linked to another polypeptide or protein or a polysaccharide. linked polypeptide or protein may comprise a surface protein or peptide corresponding thereto pathogenic bacteria, which may be the P1, P2 or P6 outer membrane protein of H. influenzae. The polysaccharide preferably comprise a PRP molecule from H. influenzae.

BRIEF DESCRIPTION OF THE FIGURES

The present invention will be further understood from the following description with reference to the drawings, in which:

Figure 1A shows the nucleotide sequence of the D15 gene from <u>H. influenzae</u> type b Ca strain (SEQ ID NO: 1) and its deduced amino acid sequence (SEQ ID NO: 2);

Figure 1B shows the nucleotide sequence of the D15 gene from <u>H. influenzae</u> type b Eagan strain (SEQ ID NO: 3) and its deduced amino acid sequence (SEQ ID NO: 4);

Figure 1C shows the nucleotide sequence of the D15 gene from <u>H. influenzae</u> type b MinnA strain (SEQ ID NO. 5) and its deduced amino acid sequence (SEQ ID NO: 6);

Figure 1D shows the nucleotide sequence of the D15 gene from <u>H. influenzae</u> non-typeable SB33 (SEQ ID NO. 7) and its deduced amino acid sequence (SEQ ID NO: 8);

Figure 1E shows the nucleotide sequence of the D15 gene from <u>H. influenzae</u> non-typeable PAK 12085 (SEQ ID NO. 9) and its deduced amino acid sequence (SEQ ID NO: 10);

Figure 1F shows an alignment of the nucleotide 35 sequences of the D15 genes (SEQ ID NOS: 1, 3, 5, 7 and 9)

10

15

30

35

obtained from different <u>H. influenzae</u> isolates (typeable, Ca, Eagan and MinnA; nontypeable SB33 and PAK 12085);

Figure 2 shows restriction maps of clones pUC19/D15 (Ca), DS-712-2-1 (Eagan), DS-691-1-5 (MinnA), JB-1042-5-1 (SB33), and JB-1042-9-4 (PAK 12085). $H = \frac{\text{HindIII}}{\text{EcoRI}}$; $R = \frac{\text{EcoRI}}{\text{EcoRI}}$; $S = \frac{\text{Sau}}{\text{Sau}}$ I; and $Xb = \frac{\text{XbaI}}{\text{Sau}}$;

Figure 3 shows an alignment of the amino acid sequences of D15 outer membrane proteins (SEQ ID NOS: 2, 4, 6, 8 and 10) obtained from different <u>H. influenzae</u> isolates (typeable, Ca, Eagan and MinnA; nontypeable, SB33 and PAK 12085). Amino acids are represented by the conventional one-letter code. The Ca D15 sequence is used as reference and the dots indicate amino acid residues which are identical to those of the Ca D15 outer membrane protein;

Figure 4 shows the construction of a plasmid (DS-880-1-2) expressing full-length SB33 D15 (rD15) from the strong inducible T7 promoter;

Figure 5 shows an SDS-PAGE analysis of native D15 20 affinity-purified from <u>H. influenzae</u> strain 30;

Figure 6 shows an SDS-PAGE analysis of sequential fractions obtained during the purification of the full-length rD15 expressed in <u>E. coli</u> containing plasmid DS-880-1-2;

Figure 7 shows guinea pig IgG antibody responses to full length rD15. The arrows indicate the immunization schedule. Bleeds were taken at 0, 2, 4, 6 and 8 weeks. The bars represent the standard deviation;

Figure 8 shows mouse IgG antibody responses to full length rD15. The arrows indicate the immunization schedule. Bleeds were taken at 0, 1, 4, 5 and 7 weeks. The bars represent the standard deviation;

Figure 9 shows an SDS-PAGE analysis of the N-terminal rD15 fragment purified from GST-(D15 fragment) fusion protein. Lanes: 1, prestained low molecular weight markers (14kDa, 21kDa, 31 kDa, 45kDa, 68kDa,

20

25

97kDa); 2, GST standard; 3, GST-(D15 fragment) fusion protein; 4, fusion protein cleaved by thrombin; 5, N-terminal rD15 fragment; 6, GST; 7, low molecular weight markers;

Figure 10 shows guinea pig IgG antibody response to N-terminal rD15 fragment. The arrows indicate the immunization schedule. Bleeds were taken at 2, 4, 6 and 8 weeks. The bars represent the standard deviation; and

Figure 11 shows the hydrophilicity plot of D15
10 established by using a window average across 7 residues
according to Hope, 1986.

GENERAL DESCRIPTION OF THE INVENTION

Any <u>Haemophilus</u> strains that have D15 genes may be conveniently used to provide the purified and isolated nucleic acid molecules (which may be in the form of DNA molecules), comprising at least a portion coding for a D15 outer membrane protein as typified by embodiments of the present invention. Such strains are generally available from clinical sources and from bacterial culture collections, such as the American Type Culture Collection. <u>H. influenzae</u> strains may include types a, b and c strains, non-typeable strains and other bacteria that produce a D15 protein, fragment or analog thereof. Appropriate strains of <u>Haemophilus</u> include:-

H. influenzae type b strain Ca;

H. influenzae type b strain MinnA;

H. influenzae type b strain Egan;

H. influenzae non-typeable b strain SB33; or

H. influenzae non-typeable b strain PAK 12085.

In this application, the term D15 outer membrane protein is used to define a family of D15 proteins which includes those having naturally occurring variations in their amino acid sequences as found in various strains of, for example, <u>Haemophilus</u>. The purified and isolated DNA molecules comprising at least a portion coding for D15 outer membrane protein of the present invention also

10

15

20

25

include those having naturally occuring variations in their nucleic acid sequences as found in various strains of, for example <u>Haemophilus</u> and those DNA molecules encoding functional analogs of D15 outer membrane protein. In this application, a first protein is a functional analog of a second protein if the first protein is immunologically related with and/or has the same function as the second protein. The functional analog may be, for example, a fragment of the protein or a substitution, addition or deletion mutant thereof.

In aspects of the present invention, the D15 gene was isolated from <u>H. influenzae</u> type b strain Ca as shown in Figure 1A; <u>H. influenzae</u> type B Eagan, Figure 1B; <u>H. influenzae</u> type b MinnA, Figure 1C; non-typeable <u>H. influenzae</u> SB33, Figure 1D; non-typeable <u>H. influenzae</u> PAK 12085, Figure 1E. A comparison of the nucleic acid sequences of the D15 genes and of the deduced amino acid sequences of the D15 outer membrane proteins from these strains of <u>H. influenzae</u> showed the genes and proteins to be highly conserved (Figures 1F and 3). The consensus sequence (SEQ ID NO: 55) for the D15 gene is shown in Figure 1F.

The purified and isolated DNA molecules comprising at least a portion coding for a D15 outer membrane protein of a species of <u>Haemophilus</u>, typified by the embodiments described herein, are advantageous as:

- nucleic acid probes for the specific identification of <u>Haemophilus</u> strains <u>in vitro</u> or <u>in vivo</u>;
- the products encoded by the DNA molecules are 30 useful as diagnostic reagents, antigens for the production of Haemophilus-specific antisera, for vaccination against the diseases caused by species of Haemophilus and detecting infection by Haemophilus; and
- peptides corresponding to portions of the D15 outer membrane protein as typified by the embodiments

10

described herein are advantageous as diagnostic reagents, antigens for the production of <u>Haemophilus</u>-specific antisera, for vaccination against the diseases caused by species of <u>Haemophilus</u> and for detecting infection by <u>Haemophilus</u>.

Reference will now be made in detail to the presently preferred embodiments of the invention, which together with the following Examples, serve to explain the principle of the invention. For clarity of disclosure, and not by way of limitation, the detailed description of the invention is divided into the following sections:

(i) The DNA sequences coding for the outer membrane protein D15 from H. influenzae type b Ca strain.

15 A clone producing the outer membrane protein designated D15 of H. influenzae type b (Hib) was isolated by screening a genomic library with H. influenzae type b OMP-specific polyclonal antibodies as previously described by Berns and Thomas 1965; Thomas and Rossi 20 The DNA fragment encoding the D15 protein was isolated, subcloned into pUC19 to produce pUC19/D15 (Figure 2) and used to transform E. coli HB101 as described in Example 1. Plasmid DNA was prepared from two individual colonies of E. coli HB101 containing the 25 pUC19/D15 plasmid. Sequencing was performed on an ABI DNA sequencer model 370A using dye-terminator chemistry and oligonucleotide primers which had been synthesized on an ABI DNA synthesizer model 380B, and purified chromatography. Nucleotide sequence analysis of the D15 30 gene revealed that it contains a putative promoter and an open reading frame encoding 789 amino acids (Figure 1A).

The first 19 amino acid residues of the translated open reading frame form a typical leader sequence as found in other <u>H. influenzae</u> type b outer membrane proteins, such as Pl and P2. The N-terminal sequence of immuno-affinity purified native D15 antigen was

determined by automated Edman degradation using the ABI 477A protein sequencer and was found to be Ala-Pro-Phe, which is identical to the N-terminal amino acid sequence Ala-Pro-Phe-Val-Ala-Lys- (SEQ ID NO: 11) predicted from an analysis of the sequence of the D15 gene presented in Figure 1A.

(ii) The sequence of D15 genes from other H. influenzae strains.

D15 genes were isolated from other H. influenzae strains by screening the chromosomal libraries of \underline{H} . 10 influenzae type b strains Eagan, Minn A and the nontypeable H. influenzae (NTHi) strains SB33 and PAK 12085, as described in Examples 2, 3 and 4. Hybridizationpositive clones were plated and submitted to a second 15 round of screening. The restriction maps of the clones are shown in Figure 2. obtained The nucleotide sequences of the D15 genes were determined for all these clones (Figures 1B to 1E) and their derived amino acid sequences compared (Figure 3). The D15 amino acid sequences of the three $\underline{H.\ influenzae}$ type b strains were 20 identical and only a few amino acid differences were observed in the amino acid sequence of the D15 protein from the non-typeable strains (Figure 3).

(iii) Expression of D15 and its fragments in E. coli.

Since D15 is expressed in small quantities by strains of <u>H. influenzae</u>, it is advantageous to either express this antigen as a recombinant protein in a heterologous system, such as <u>E. coli</u>, or to modify the <u>H. influenzae</u> organism to enhance native D15 expression. The <u>Hind III/Eco</u> RI fragment of <u>H. influenzae</u> type b Ca strain DNA encoding the full length D15 protein was expressed in pUC19 but not pUC18, suggesting that the <u>lac</u> promoter is helping to express the D15 gene in <u>E. coli</u>, even though the native D15 gene promoter is present. The T7 expression system is a tightly controlled, inducible system which has great utility in expression of

heterologous proteins in E. coli. The T7 expression system is described in U.S. Patent 4,952,496. were, therefore, constructed which utilize the T7 system to express a mature D15 protein that contains additional methionine residue at the amino terminus. The 5 D15 signal sequence was removed during this construction process. A full length recombinant D15 (termed rD15) was expressed in inclusion bodies which allow the D15 protein to be readily purified. The D15 genes from H. influenzae type b strain Ca and <u>H. influenzae</u> non-typeable SB33 10 strain have been expressed at high levels in E. coli using the T7 system to permit production of large quantities of rD15 protein. The construction of clone DS-880-1-2 which expresses the SB33 D15 gene is described herein (see Figure 4 and Example 5). 15 The rD15 protein was immunologically similar to its native counterpart isolated from <u>H. influenzae</u> typeable and non-typeable strains (see below). Thus, rD15 may be used as a crossreactive antigen in a diagnostic kit to detect many, if not all, strains of H. influenzae and other bacteria that produce a D15 outer membrane protein or analog thereof. Alternatively, rD15 can be used as an antigen to specifically detect the presence of H. influenzae in a sample.

25 A truncated D15 fragment was expressed in E. coli as a fusion protein with glutathione S-transferase (GST), as described in Example 6. The construction was designed to express the N-terminal fragment of the D15 protein. fusion protein was expressed at high levels from a pGEX-2T construction and the N-terminal fragment was cleaved 30 from the GST carrier protein by treatment with thrombin. This procedure generated a molecule termed the N-terminal rD15 fragment which encompasses amino acids 63-223 of the D15 protein. This N-terminal rD15 fragment was highly 35 immunogenic and elicited protective antibodies against challenge with live H. influenzae.

(iv) Purification of native D15 from <u>H. influenzae</u> cell paste.

The present invention also provides a method to prepare purified native D15 protein from H. influenzae. The protein is extracted and affinity-purified from the cell pastes of either <u>H. influenzae</u> typeable or nonisolates by a procedure involving dissolution of the protein in an aqueous detergent solution (see Example 13). The native D15 protein from a 10 non-typeable H. influenzae strain 30 was solubilized with a 50 mM Tris-HCl/ 0.5% Triton X-100/ 10 mM EDTA buffer, pH 8.0 and further purified on a D15-specific monoclonal antibody affinity column (Figure 5A). An 80 kDa protein was eluted from the column with 50 mM diethylamine, pH 12.0 and shown to react with a D15-specific monoclonal 15 antibody on immunoblot analysis (Fig. 5B). The native D15 is also highly immunogenic in experimental animals. Rabbit anti-D15 antisera reacted with all H. influenzae isolates as determined by immunoblot analyses.

20 (v) Purification of a full-length recombinant D15 protein expressed in E. coli.

A full-length recombinant D15 (rD15) protein was expressed in inclusion bodies in <u>E. coli</u>. As shown in Figure 6, purification of rD15 inclusion bodies was achieved by a sequential extraction of the <u>E. coli</u> cell lysate with 50 mM Tris-HCl, pH 8.0, then 50 mM Tris containing 0.5% Triton X-100 and 10 mM EDTA, pH 8.0. After centrifugation, more than 95% of the proteins in the resulting pellet was an 80 kDa protein by SDS-PAGE analysis, that reacted with a D15-specific monoclonal antibody on an immunoblot. The N-terminal sequence of the rD15 was found to be Met-Ala-Pro-Phe-Val-Lys-Asp-(SEQ ID NO: 54) which is identical to the predicted amino acid sequence.

The rD15 inclusion bodies were solubilized with a mixture of PBS, 0.5% Triton X-100, 10 mM EDTA and 8 M

25

15

20

25

35

urea (see Example 8). After dialysis against PBS to remove urea, more than 80% of the D15 protein remained soluble. This soluble rD15 antigen was used for the immunogenicity studies described below. From shake-flask experiments, it was estimated that about 10 mg of soluble rD15 protein was obtained from 1 L of E. coli bacterial culture. It is clear that growing the recombinant E. coli strains under optimised fermentation conditions significantly increase the level of rD15 production.

10 (vi) Immunogenicity of the full-length recombinant D15 protien (rD15).

The immunogenicity of the full-length rD15 protein was studied in guinea pigs and mice. Using the immunization protocols described in Figure 7, a 15 μ g dose of rD15 induced high IgG titers in guinea pigs when administered in the presence of either Freund's adjuvant or AlPO₄. In the mouse dose-response study, the protein appeared to be immunogenic at a dose as low as 5 μ g in either Freund's adjuvant (Figure 8A) or AlPO₄ (Figure 8B).

The protective ability of rD15 against <u>H. influenzae</u> type b infection was examined in the infant rat model of bacteremia essentially as described by Loeb (1987). Thus, infant rats passively immunized with guinea pig anti-rD15 antisera were significantly less bacteremic than controls injected with pre-bleed sera, which is consistent with the previous report by Thomas et al. (1990).

(vii) Purification and characterization of the N-30 terminal rD15 fragment.

The truncated rD15 fragment corresponding to the N-terminus of the D15 protein (residues 22 to 223) as described in Example 6, was expressed in <u>E. coli</u> as a soluble protein fused to GST. The fusion protein (46 kDa) was readily extracted using phosphate buffered saline (PBS). Purification of the GST-D15 fragment fusion

protein was achieved by a single-step affinity purification process on a glutathione-Sepharose 4B column (Figure 9, Lane 3). Cleavage of the 46 kDa fusion protein with thrombin yielded two fragments (Figure 9, Lane 4), a 26 kDa protein which corresponded to a 5 purified GST standard (Figure 9, Lane 2), and a 20 kDa polypeptide which had the size expected for the Nterminal rD15 fragment (amino acid residues 63 to 223), respectively. Separation of these two proteins was achieved by a second round of glutathione-Sepharose 4B 10 affinity chromatography. From shake-flask experiments, it was estimated that about 1 mg of purified N-terminal rD15 fragment was recovered from 1 L of E. coli bacterial It is clear that growing the recombinant E. culture. coli strains under optimised fermentation conditions will significantly increase the level of N-terminal rD15 fragment production.

The identity of the 20 kDa polypeptide and the 26 kDa protein was confirmed by both immunoblotting and 20 protein sequencing. The N-terminal sequence of the 20 kDa polypeptide was found to be NH2-Ser-Leu-Phe-Val-Ser-Gly-Arg-Phe-Asp-Asp-Val-Lys-Ala-His-Gln-Glu-Gly-Asp-Val-Leu-Val-Val-Ser- (SEQ ID NO: 12), which corresponds to residues 63 to 85 of the primary sequence of D15. This result indicates that there is a spurious thrombin 25 cleavage site within the D15 sequence and that the first 42 amino acids of the rD15 fragment are cleaved off during thrombin digestion. Thus, the final N-terminal rD15 fragment was 161 amino acids in length corresponding 30 to residues 63 to 223 of the primary sequence of D15. The N-terminal sequence obtained for the 26 kDa protein $(NH_2-$ Met-Ser-Pro-Ile-Leu-Gly-Tyr-Trp-Lys- - SEQ ID NO: 13) confirmed that it was GST.

(viii) Immunogenicity of the N-terminal rD15 fragment.

The immunogenicity of the N-terminal rD15 fragment was tested in guinea pigs using various adjuvants. Using

15

15

20

25

30

the immunization protocols described in Figure 10, a 10 μ g dose of N-terminal rD15 fragment induced a good booster response in guinea pigs with almost all the adjuvants tested. The highest anti-D15 IgG titer was observed in the group of guinea pigs immunized with N-terminal rD15 fragment in Freund's adjuvant. The second best adjuvant was Titermax (CytRx Inc.). The other two adjuvants, TPAD4 (tripalmityl-Cys-Ser-Glu₄) and AlPO₄ were equally potent.

10 (ix) Protective ability of the N-terminal rD15 fragment against <u>H. influenzae</u> type b challenge.

An <u>in vivo</u> challenge model for a assessing the protective abilities of antigen against diseases caused by <u>Haemophilus</u> is the infant rat model of bacteremia as described by Loeb 1987. The protective ability of the N-terminal rD15 fragment against <u>H. influenzae</u> type b challenge was examined in this rat model. As illustrated in Table 1, infant rats passively immunized with rabbit anti-N-terminal rD15 fragment antisera showed significantly lower bacteremia compared to those injected with pre-bleed sera.

Since passively transferred antisera against the N-terminal rD15 fragment were found to be protective in the infant rat model of bacteremia, it was of interest to identify the protective epitope(s) of this N-terminal rD15 fragment. The first nine overlapping peptides of the D15 protein as listed in Table 2 were chemically synthesized based upon the amino acid sequence derived from the sequence of the D15 gene from H. influenzae type b Ca (Figure 1). These synthetic peptides were assessed for their reactivities with either rabbit or guinea pig antisera raised against purified N-terminal rD15 fragment by ELISAs. As shown in Table 3, both guinea pig and rabbit antisera reacted with a cluster of D15 peptides, including peptides D15-P4 to D15-P8 encompassing residues 93 to 209 of the D15 primary sequence.

30

35

Further studies were performed to determine whether the protection against <u>H. influenzae</u> type b observed using rabbit anti-D15 antisera in infant rats could be neutralized by D15 peptides. In the first experiment, a rabbit anti-N-terminal rD15 fragment antiserum was injected into a group of seven infant rats in the presence or absence of a mixture of the nine D15 peptides (D15-P2 to D15-P10). Animals in the positive control group were injected with the rabbit anti-N-terminal rD15 fragment antiserum mixed with purified D15 fragment and 10 the negative control group was injected with a mixture of the nine peptides only. As illustrated in Table 4, infant rats passively immunized with a rabbit anti-N-terminal rD15 fragment antiserum (group #1) showed a significantly lower bacteremia level (3%, $p = 1.2x10^{-7}$) compared to 15 those in the negative control group (group #4, 100%), which was consistent with the previously obtained results. The protection mediated by the rabbit anti-Nterminal rD15 fragment antiserum was largely neutralized 20 by the addition of purified N-terminal rD15 fragment (group #3, 64%), as indicated by the lack of significant difference in the bacteremia level between group #3 and group #4 (p = 0.09). Although the addition of the mixture of nine D15 peptides only slightly neutralized the protection conferred by the antiserum (group #2, 13%) as compared to group #1 (3%), the difference in bacteria counts between these two groups was statistically significant (p = 0.0037).

To more clearly define the protective epitope(s) of the N-terminal rD15 fragment, the above experiment was repeated with a mixture of five peptides (peptides D15-P4 to D15-P8) which were chosen for their strong reactivities with the rabbit anti-N-terminal fragment antiserum. The results obtained from this second experiment showed that the protection observed using rabbit anti-N-terminal rD15 fragment (Table 5, group #1)

15

was completely blocked by the addition of this mixture of five peptides (Table 5, group #2, 106%, $p = 0.53 \times 10^{-8}$). These results strongly indicate that a cocktail of D15 synthetic peptides may be used as immunogens to induce protective antibodies against <u>H. influenzae</u>.

(x) Epitope prediction and peptide synthesis.

To map the immunodominant T-cell or B-cell epitopes of D15, overlapping synthetic peptides covering the entire D15 protein sequence (Table 2 - SEQ ID NO: 14 to 49) were synthesized using the t-Boc solid-phase peptide synthesis as described in Example 15. The peptides were chosen based on their high index of hydrophillic β -turns estimated by secondary structure prediction analysis (Figure 11). Such peptides are likely to be surface-exposed and antigenic. Peptides more than 25 residues in length were selected to better mimic native epitopes.

(xi) Identification and characterization of immunodominant epitopes of D15 using synthetic peptides.

the linear B-cell epitopes of overlapping synthetic peptides representing the entire 20 sequence of D15 were individually coated onto ELISA plates and probed with several anti-rD15 antisera as described in Example 19. The results are summarized in Table 6. Mouse antisera raised against rD15 reacted with 25 all D15 peptides, but the major epitopes were located within peptides D15-P8 (residues 180-209 - SEQ ID NO: 21), D15-P10 (residues 219-249 - SEQ ID NO: 23), D15-P11 (residues 241-270 - SEQ ID NO: 24), and D15-P26 (residues 554-582 - SEQ ID NO: 39), respectively. Rabbit anti-D15 30 antisera recognized only peptides D15-P4 (residues 93-122 - SEQ ID NO: 17), D15-P14 (residues 304-333 - SEQ ID NO: 27) and D15-P36 (residues 769-798 - SEQ ID NO: 49). Guinea pig antisera raised against rD15 reacted with peptides D15-P2 (residues 45-72 - SEQ ID NO: 15), D15-P4 35 (residues 93-122 - SEQ ID NO: 17), D15-P6 (residues 135-164 - SEQ ID NO: 19), D15-P8 (residues 180-209 - SEQ ID

NO: 21), D15-P14 (residues 304-333 - SEQ ID NO: 27), D15-(residues 577-602 - SEQ ID NO: 40). immunodominant linear B-cell epitopes of D15 were thus found to be located within peptides D15-P4 (residues 93-122 - SEQ ID NO: 17) and D15-P14 (residues 304-333 - SEQ 5 ID NO: 27), since these are the only two peptides recognized by rD15-specific antisera from all three animal species. These results indicate that the peptides containing the linear B-cell epitope sequences described above can be used as target antigens in, for example, 10 diagnostic kits to detect the presence of anti-D15 and anti-H. influenzae antibodies in samples.

(xii) Identification and characterization of immunodominant T-cell epitopes of D15 using synthetic peptides.

The importance of cytokine networks in the immune and inflammatory responses in immunity and inflammation and their alteration in pathology is becoming more evident as new members of the cytokine family are identified and characterized. Mills et al. (1993) have 20 recently reported that there is a rapid clearance of B. pertussis from the lungs of mice on challenge six weeks respiratory infection or following immunizations with the whole-cell pertussis vaccine. Spleen cells from these immunized mice were found to 25 secrete high levels of IL-2 and IFN- γ and low levels of IL-5 in the presence of pertussis antigen (pertussis toxoid, filamentous haemagglutinin (FHA) and pertactin). This result suggests that Th1 cell (T-cells producing high levels of IL-2 and IFN- γ) proliferation is very 30 important for recovering from respiratory infection. The generation of Th1 and Th2 cell subsets is regulated by the balance between different groups of cytokines, predominantly IL-12 and IL-4 (Trinchieri, 1993). IL-12 35 IL-4 are responsible for Th1 and Th2 cells differentiation, respectively. One of the roles of Th2

cells in the immune system is to provide helper activity for eliciting high levels of antigen-specific antibodies following immunization. Antigens containing Th1 epitope(s) stimulate antigen-specific T-cells to produce high levels of IL-2 and IFN- γ , whereas Th2 epitope(s) induce high levels of IL-4 expression. Th0 epitope(s) stimulate the synthesis of IFN- γ and IL-4.

Little is known about the cellular immune response to outer membrane proteins of H. influenzae and its role 10 in the protection against H. influenzae infection and To this end, the inventors performed studies of the cellular response elicited in mice following rD15 immunization. D15-specific T-cell epitopes determined using D15 peptides and T-cell lines obtained from five BALB/c mice immunized with rD15 (see Example 15 The lymphocyte proliferative responses of the D15specific T-cell lines to overlapping D15 peptides were determined in conventional cytokine assays as described in Example 24. The results summarized in Table 7, revealed that stimulation only with certain synthetic 20 peptides elicited proliferative responses and the release of specific cytokines. Synthetic peptides corresponding to residues 114-143 (D15-P5 - SEQ ID NO: 18), 282-312 (D15-P13 - SEQ ID NO: 26) and 577-602 (D15-P27 - SEQ ID NO: 40), and 219-249 (D15-P10 - SEQ ID NO: 23), 262-291 25 (D15-P12 - SEQ ID NO: 25), 390-416 (D15-P18 - SEQ ID NO: 31), 410-435 (D15-P19 - SEQ ID NO: 32) 554-582 (D15-P26 -SEQ ID NO: 39), 596-625 (D15-P28 - SEQ ID NO: 41), 725-750 (D15-P34 - SEQ ID NO: 47) and 745-771 (D15-P35 - SEQ ID NO: 48) were shown to be highly stimulatory for rD15-30 specific BALB/c Th0 cells and Th1 cells, respectively. Therefore, these immunodominant T-cell epitopes can be used as autologous carriers for PRP, and/or OMP B-cell epitopes to enhance their immunogenicity. The Th1 cell 35 epitopes identified above may be useful in the H.

<u>influenzae</u> vaccine formulations to induce <u>H. influenzae</u>specific cellular immune responses.

(xiii) Immunogenicity of D15 peptides.

To determine whether synthetic D15 peptides were immunogenic free peptides were assessed individually for 5 their immunogenicity. Rabbit and guinea pig anti-peptide antisera were tested for their reactivities with the immunizing peptides as well as with native D15 and rD15 by ELISA and immunoblotting. As shown in Table 8, all 10 guinea pig anti-D15 peptide antisera except those raised against D15-P26 (SEQ ID NO: 39), D15-P29 (SEQ ID NO: 42), D15-P30 (SEQ ID NO: 43) and D15-P31 (SEQ ID NO: 44) were shown to be immunogenic by ELISAs. The induction of high titers of peptide-specific IgG antibodies by 15 peptides clearly indicates that most peptides contain both a functional T-helper determinant and a B-cell In addition, these anti-peptide antisera epitope(s). recognised D15 in the immunoblot assay. Since most peptides contain potent functional T-helper determinant(s) and induce strong IgG antibody responses 20 in mammals, they are candidate immunogens for inclusion in an <u>H. influenzae</u> vaccine preparation. D15 peptidespecific antisera cross-reacted with D15 from nontypeable strains of H. influenzae judged by as 25 immunoblotting. This finding indicates that immunogenic D15 peptides contain epitopes which are highly conserved among typeable and non-typeable strains of H. influenzae. In addition, polyclonal antibodies against these epitopes are useful to detect <u>H. influenzae</u> in biological samples. 30

Therefore, these conserved epitopes of D15 can be used either individually or in combination to prepare cross-reactive synthetic immunogens against typeable and non-typeable strains of <u>H. influenzae</u> and other bacteria that produce D15 protein, a fragment or an analog thereof. Peptides described above can be further polymerized, or modified with lipids as lipopeptides, or

15

20

25

30

linked to polysaccharides including PRP as synthetic glycopeptide or lipoglycopeptide conjugates to produce alternate vaccines. These vaccines can be used to immunize against diseases caused by H. influenzae when administered to mammals, for example, by intramuscular or parenteral route, or when delivered using microparticles, capsules, liposomes and targeting molecules, such as toxins or fragments thereof, and antibodies, to cells of the immune system or mucosal surfaces.

(xiv) Utility of D15 as carrier protein for the production of glycoconjugates.

To determine whether D15 may serve both as a protective antigen and a carrier, D15-PRP conjugation experiments were performed as described in Example 14. The D15-PRP conjugates were found to be highly immunogenic in rabbits and able to elicit both anti-D15 and anti-PRP IgG antibody responses as judged by D15-specific ELISA and PRP-BSA immunoassay (Table 9). These results clearly demonstrate the practical utility of D15 as a carrier protein for glycoconjugation technology.

In preferred embodiments of the present invention, the carrier function of D15 can be generally utilized to prepare chimeric molecules and conjugate vaccines against pathogenic bacteria, including encapsulated bacteria. Thus, the glycoconjugates of the present inventions may be applied to vaccinations to confer protection against infection with any bacteria having polysaccharide antigens, including, for example, Haemophilus influenzae, Streptococcus pneumoniae, Escherichia coli, Neisseria meningitidis, Salmonella typhi, Streptococcus mutans, Cryptococcus neoformans, Klebsiella, Staphylococcus aureus and Pseudomonas aeruginosa.

In another embodiment, the carrier function of D15
may be used, for example, to induce immunity toward
abnormal polysaccharides of tumor cells, or to produce

anti-tumor antibodies that can be conjugated to chemotherapeutic or bioactive agents.

Accordingly, the present invention provides the primary sequence and the preparation of an antigen (D15) of <u>H. influenzae</u> that can be used in the prevention and diagnosis of diseases caused by Haemophilus. particular, the inventors discovered that recombinant D15 fragments, can elicit protective antibody responses against live <u>H. influenzae</u> type b bacteria 10 challenge. Thus, the present inventions have utility in vaccines. The invention also discloses the nucleotide sequences of the D15 genes isolated from both H. influenzae type b strains and non-typeable isolates. The DNA segments encoding D15 are disclosed and show minor 15 polymorphism in both their nucleotide and derived amino acid sequences (Figures 1F and 3). These DNA segments may be used to provide an immunogen essentially free from other <u>H. influenzae</u> antigens (such as lipooligosaccharides (LOS)) through the application of 20 recombinant DNA technology. The present disclosure further provides novel techniques which can be employed for preparing essentially pure D15 or fragments thereof. as well as functional analogs. The recombinant D15 protein, fragment or analog thereof, may be produced in 25 suitable expression system, such as E. coli, <u>Haemophilus</u> Bordetella, Bacillus, Fungi, Baculovirus, Poxvirus, vaccinia or mammalian expression systems.

In one embodiment, the present invention concerns
the process of preparing vaccine compositions which
include purified recombinant D15 protein (rD15) or rD15
fragments that are immunologically cross-reactive with
native D15. In particular, the gene coding the entire
D15 protein and a DNA segment encoding an N-terminal rD15
fragment fused to the glutathione-S-transferase gene have
been constructed and expressed in <u>E. coli</u>. The expressed

10

15

20

25

30

35

rD15 protein and its fragments were found to cross-react immunologically with the native D15 antigen isolated from both typeable and non-typeable H. influenzae isolates and thus represent cross-reactive immunogens for inclusion in a vaccine against diseases caused by H. influenzae. Furthermore, Haemophilus convalescent serum recognized D15 purified from H. influenzae as described herein, rD15 and N-terminal rD15 fragment.

another embodiment, the present invention provides a gene coding for the outer membrane protein D15 from <u>H. influenzae</u> having the specific nucleotide sequences described herein or ones substantially homologous thereto (i.e. those which hybridize under stringent conditions to such sequences), for genetically engineering hybrids or chimeric proteins containing a D15 fragment fused to another polypeptide or protein or a polysaccharide, such as <u>H. influenzae</u> outer membrane proteins, for example, P1, P2, or P6 or PRP. result, the hybrids, chimeric proteins or glycoconjugates may have higher protectivity against H. influenzae than D15, or P1, or P2, or P6, or PRP alone.

Thus, D15 outer membrane protein can function both as a protective antigen and as a carrier in a conjugate vaccine to provide autologous T-cell priming, wherein the hapten part of the conjugate is the capsular polysaccharide moiety (PRP) of H. influenzae. This D15carbohydrate conjugate can elicit antibodies against both PRP and D15, and thus should enhance the level of protection against H. influenzae-related diseases, especially in infants.

In another embodiment, the present invention comprises an essentially pure form of at least one protein or peptide containing an amino acid sequence corresponding to at least one antigenic determinant of D15, which peptide is capable of eliciting polyclonal antibodies against <u>H. influenzae</u> in mammals. These D15-

specific antibodies are useful in test kits for detecting the presence of <u>H. influenzae</u> in biological samples. peptides can have, for example, the amino acid sequences corresponding to residues 20-49, 45-74, 68-99, 93-122, 5 114-143, 135-164, 157-187, 180-209, 199-228, 219-249, 241-270, 262-291, 282-312, 304-333, 325-354, 367-396, 390-416, 410-435, 430-455, 450-477, 471-497, 491-516, 511-538, 532-559, 554-582, 577-602, 596-625, 619-646, 641-666, 662-688, 681-709, 705-731, 725-750, 10 745-771, 769-798 (SEQ ID NOS: 14 to 49) of the D15 protein of the <u>H. influenzae</u> type b Ca strain, respectively, as set forth in Table 2 below, or any portion, variant or mutant thereof which retains immunogenicity.

15 In yet another embodiment, the present invention provides pure native D15 protein, extracted and chromatographically purified cultures from Η. influenzae typeable or non-typeable isolates. The novel procedures involves extraction of the D15 protein from cell paste by techniques known for other outer membrane 20 proteins, with an aqueous detergent solution, followed by purification by centrifugation and chromatography. purified native D15 antigen can be used to immunize mammals against diseases caused by H. influenzae, for 25 example, by the intramuscular or the parenteral routes, or by delivering it using microparticles, capsules, liposomes and targeting molecules, such as toxins or fragments thereof, and antibodies.

Another aspect of the present invention is that the

30 D15 outer membrane protein, fragments or analogs thereof
or peptides corresponding to portions of D15 may be
components of a multivalent vaccine against otitis media.
This multivalent vaccine comprises at least one
immunogenic determinant of D15 as described herein, along

35 with at least one protective antigen isolated from
Streptococcus pneumoniae, Branhamella (Moroxella)

20

25

<u>catarrhalis</u>, <u>Staphylococcus</u> <u>aureus</u>, or respiratory syncytial virus, in the presence or absence of adjuvant.

The D15 peptides (Table 2) or any portion, variant or mutant thereof, can easily be synthesized either manually or with a commercially available peptide synthesizer, such as the Applied Biosystems Model 430A synthesizer.

It is clearly apparent to one skilled in the art, that the various embodiments of the present invention have many applications in the fields of vaccination, diagnosis, and treatment of diseases <u>Haemophilus</u> infections, and the generation of immunological reagents. A further non-limiting discussion of such uses is further presented below.

15 1. Vaccine preparation and use

Immunogenic compositions, suitable for use as vaccines, may be prepared from immunogenic D15 outer membrane protein, fragments or analogs thereof and/or peptides corresponding to portions of D15 as disclosed herein. The vaccine elicits an immune response which produces antibodies, including anti-D15 outer membrane protein antibodies and antibodies against D15 that are opsonizing or bactericidal. Should the vaccinated subject be challenged by Haemophilus, the antibodies bind to the D15 outer membrane protein and thereby inactivate the bacterium. Opsonizing and bactericidal antibodies represent examples of antibodies useful in protection against disease.

Vaccines containing peptides are generally well
known in the art, as exemplified by U.S. Patents
4,601,903; 4,599,231; 4,599,230; and 4,596,792; all of
which references are incorporated herein by reference.
As to any further reference to patents and references in
this description, they are as well hereby incorporated by
reference without any further notice to that effect.
Vaccines may be prepared as injectables, as liquid

solutions or emulsions. The D15 outer membrane protein, fragments or analogs thereof or peptides corresponding to portions of D15 may be mixed with physiologicallyacceptable excipients which are compatible with the D15 outer membrane protein, fragments, analogs or peptides. may Excipients include, water, saline, glycerol, ethanol, and combinations thereof. The vaccine may further contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents, or adjuvants to enhance the effectiveness of the vaccines. Methods of achieving adjuvant effect for the vaccine includes use of agents, such as aluminum hydroxide or phosphate (alum), commonly used as 0.05 to 0.1 percent solution in phosphate buffered saline. Vaccines may be administered parenterally, by injection subcutaneously or intramuscularly. Alternatively, other modes of administration including suppositories and oral formulations may be desirable. For suppositories, binders and carriers include, may for example, polyalkalene glycols or triglycerides. Oral formulations may include normally employed incipients such as, for example, pharmaceutical grades of saccharine, cellulose, magnesium carbonate and the like. These compositions take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and contain 10-95% of the D15 outer membrane protein, fragment analogs and/or peptides.

The vaccines are administered in a manner compatible with the dosage formulation, and in an amount which is therapeutically effective, protective and immunogenic. The quantity to be administered depends on the subject to be treated, including, for example, the capacity of the individual's immune system to synthesize antibodies, and if needed, to produce a cell-mediated immune response. 35 Precise amounts of active ingredient required to be administered depend on the judgment of the practitioner.

5

10

15

20

25

10

15

20

25

30

35

However, suitable dosage ranges are readily determinable by one skilled in the art and may be of the order of micrograms of the D15 outer membrane protein, analog, fragment and/or peptides. Suitable regimes for initial administration and booster doses are also variable, but may include an initial administration followed by subsequent administrations. The dosage of the vaccine may also depend on the route of administration and varies according to the size of the host.

The nucleic acid molecules encoding the D15 outer membrane protein of the present invention may also be used directly for immunization by administration of the DNA directly, for example, by injection for genetic immunization or by constructing a live vector, such as Salmonella, BCG, adenovirus, poxvirus or vaccinia. A discussion of some live vectors that have been used to carry heterologous antigens to the immune system are discussed in, for example, O'Hagan (1992). Processes for the direct injection of DNA into test subjects for genetic immunization are described in, for example, Ulman et al. (1993).

The use of peptides in vivo may first require their chemical modification since the peptides themselves may not have a sufficiently long serum and/or tissue halflife. Such chemically modified peptides are referred to herein as peptide analogs. The term peptide analog extends to any functional chemical equivalent of a peptide characterized by its increased stability and/or efficacy in vivo or in vitro in respect of the practice of the invention. The term peptide analog is also used herein to extend to any amino acid derivative of the peptides as described herein. Peptide contemplated herein are produced by procedures that include, but are not limited to, modifications to side chains, incorporation of unnatural amino acids and/or their derivatives during peptide synthesis and the use of

20

25

30

cross-linkers and other methods which impose conformational constraint on the peptides or their analogs.

Examples of side chain modifications contemplated by
the present invention include modification of amino
groups, such as by reductive alkylation by reaction with
an aldehyde followed by reduction with NaBH₄; amidation
with methylacetimidate; acetylation with acetic
anhydride; carbamylation of amino groups with cyanate;
trinitrobenzylation of amino groups with 2, 4, 6,
trinitrobenzene sulfonic acid (TNBS); alkylation of amino
groups with succinic anhydride and tetrahydrophthalic
anhydride; and pyridoxylation of lysine with pyridoxa-5'phosphate followed by reduction with NaBH₄.

The guanidino group of arginine residues may be modified by the formation of heterocyclic condensation products with reagents such as 2, 3-butanedione, phenylglyoxal and glyoxal.

The carboxyl group may be modified by carbodiimide activation via o-acylisourea formation followed by subsequent derivatisation, for example, to a corresponding amide.

Sulfhydryl groups may be modified by methods, such carboxymethylation with iodoacetic iodoacetamide; performic acid oxidation to cysteic acid; formation of mixed disulphides with other compounds; reaction with maleimide; maleic anhydride or other substituted maleimide; formation of mercurial derivatives using 4-chloromercuribenzoate, chloromercuriphenylsulfonic acid, phenylmercury chloride, 2-chloromercuric-4-nitrophenol and other mercurials; carbamylation with cyanate at alkaline pH.

Tryptophan residues may be modified by, for example, oxidation with N-bromosuccinimide or alkylation of the indole ring with 2-hydroxy-5-nitrobenzyl bromide or sulphonyl halides. Tryosine residues may be altered by

10

30

35

2.

Immunoassays

nitration with tetranitromethane to form a 3-nitrotyrosine derivative.

Modification of the imidazole ring of a histidine residue may be accomplished by alkylation with iodoacetic acid derivatives or N-carbethoxylation with diethylpyrocarbonate.

Examples of incorporating unnatural amino acids and derivatives during peptide synthesis include, but are not limited to, use of norleucine, 4-amino butyric acid, 4-amino-3-hydroxy-5-phenylpentanoic acid, 6-aminohexanoic acid, t-butylglycine, norvaline, phenylglycine, ornithine, sarcosine, 4-amino-3-hydroxy-6-methylheptanoic acid, 2-thienyl alanine and/or D-isomers of amino acids.

The D15 outer membrane protein, analog, fragment 15 and/or peptides of the present invention are useful as antigens in immunoassays, including enzyme-linked immunosorbent assays (ELISA), RIAs and other non-enzyme linked antibody binding assays or procedures known to the art for the detection of anti-bacterial, <u>Haemophilus</u>, D15 20 and/or peptide antibodies. In ELISA assays, the D15 outer membrane protein, fragment or analogs thereof and/or peptides corresponding to portions of D15 outer membrane protein are immobilized onto a selected surface, 25 for example, a surface exhibiting a protein affinity,

such as the wells of a polystyrene microtiter plate. After washing to remove incompletely adsorbed D15 outer membrane protein, analog, fragment and/or peptides, a nonspecific protein, such as bovine serum albumin (BSA) or casein, that is known to be antigenically neutral with regard to the test sample may be bound to the selected surface. This allows for blocking of nonspecific adsorption sites on the immobilizing surface and thus decreases the background caused by nonspecific bindings of antisera onto the surface. Normally, the peptides

10

employed herein are in the range of 12 residues and up and preferably 14 to 30 residues.

The immobilizing surface is then contacted with a sample such as clinical or biological materials to be tested to а manner conducive immune complex (antigen/antibody) formation. This may include diluting the sample with diluents, such as BSA, bovine gamma globulin (BGG) and/or phosphate buffered saline (PBS)/Tween. The sample is then allowed to incubate for from 2 to 4 hours, at temperatures, such as of the order of 25° to 37°C. Following incubation, the samplecontacted surface is washed to remove non-immunocomplexed material. The washing procedure may include washing with a solution such as PBS/Tween, or a borate buffer.

Following formation of specific immunocomplexes 15 between the test sample and the bound D15 outer membrane protein, analog, fragment and/or peptides, and subsequent washing, the occurrence, and even immunocomplex formation may be determined by subjecting 20 the immunocomplex to a second antibody having specificity for the first antibody. If the test sample is of human origin, the second antibody is an antibody having specificity for human immunoglobulins and, in general, IgG. To provide detecting means, the second antibody may 25 have an associated activity, such as an enzymatic activity that will generate, for example, a color development upon incubating with an appropriate chromogenic substrate. Quantification may then achieved by measuring the degree of color generation using, for example, a visible spectra spectrophotometer. 30

3. Use of sequences as hybridization probes

The nucleotide sequences of the present invention, comprising the sequence of the D15 outer membrane protein, now allow for the identification and cloning of the D15 outer membrane protein genes from any species of

10

15

20

25

30

35

<u>Haemophilus</u> and other bacteria that have genes encoding D15 outer membrane proteins.

The nucleotide sequences comprising the sequence encoding the D15 outer membrane protein of the present invention are useful for their ability to selectively form duplex molecules with complementary stretches of other D15 genes. Depending on the application, a variety of hybridization conditions may be employed to achieve varying degrees of selectivity of the probe toward the For a high degree of selectivity, other D15 genes. stringent conditions are used to form the duplexes, such as low salt and/or high temperature conditions, such as provided by 0.02 M to 0.15 M NaCl at temperatures of between about 50°C to 70°C. For some applications, less stringent hybridization conditions are required such as 0.15 M to 0.9 M salt, at temperatures ranging from between about 20°C to 55°C. Hybridization conditions can also be rendered more stringent by the addition of increasing amounts of formamide, to destabilize the hybrid duplex. Thus, particular hybridization conditions can be readily manipulated, and will generally be a method of choice depending on the desired results.

In a clinical diagnostic embodiment, the nucleic acid sequences of the D15 outer membrane protein genes of the present invention may be used in combination with an appropriate means, such as a label, for determining hybridization. A wide variety of appropriate indicator means are known in the art, including radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of providing a detectable signal. In some diagnostic embodiments, an enzyme tag, such as urease, alkaline phosphatase or peroxidase, instead of a radioactive tag may be used. In the case of enzyme tags, colorimetric indicator substrates are known which can be employed to provide means visible to the human eye or

spectrophotometrically, to identify specific hybridization with samples containing D15 gene sequences.

The nucleic acid sequences of D15 genes of the present invention are useful as hybridization probes in solution hybridizations and in embodiments employing solid-phase procedures. In embodiments involving solid phase procedures, the test DNA (or RNA) from samples, such as clinical samples, including exudates, body fluids (e. g., serum, amniotic fluid, middle ear effusion, sputum, bronchoalveolar lavage fluid) or even tissues, is 10 adsorbed or otherwise affixed to a selected matrix or surface. The fixed, single-stranded nucleic acid is then subjected to specific hybridization with selected probes comprising the nucleic acid sequences of the D15 genes or fragments thereof of the present invention under desired 15 The selected conditions will depend on the conditions. particular circumstances based on the particular criteria required depending on, for example, on the G+C contents, type of target nucleic acid, source of nucleic acid, size 20 of hybridization probe etc. Following washing of the hybridization surface so as to remove non-specifically bound probe molecules, specific hybridization detected, or even quantified, by means of the label. selected probe should be at least 18 bp and may be in the 25 range of 30 bp to 90 bp long.

Plasmid vectors containing replicon and control sequences which are derived from species compatible with the host cell may be used for the expression of the D15 outer membrane protein genes in expression systems. The vector ordinarily carries a replication site, as well as marking sequences which are capable of providing phenotypic selection in transformed cells. For example, E. coli may be transformed using pBR322 which contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells.

30

10

15

20

The pBR322 plasmid, or other microbial plasmid or phage must also contain, or be modified to contain, promoters which can be used by the microbial organism for expression of its own proteins.

In addition, phage vectors containing replicon and control sequences that are compatible with the host microorganism can be used as a transforming vector in connection with these hosts. For example, the phage in lambda GEM^{TM} -11 may be utilized in making recombinant phage vectors which can be used to transform host cells, such as E. coli LE392.

Promoters commonly used in recombinant construction include the β -lactamase (penicillinase) and lactose promoter systems and other microbial promoters, such as the T7 promoter system. Details concerning the nucleotide sequences of promoters are known, enabling a skilled worker to ligate them functionally with plasmid The particular promoter used generally is a matter of choice depending upon the desired results. Hosts that are appropriate for expression of the transferrin receptor genes, fragment analogs or variants E. coli, thereof include Bacillus, Haemophilus, Bordetella, fungi, yeast, or the baculovirus and poxvirus expression systems may be used.

In accordance with an aspect of this invention, it is preferred to make the D15 outer membrane protein, fragment or analog thereof by recombinant methods, particularly since the naturally occurring D15 protein as purified from culture of a species of Haemophilus may include undesired contaminants, including trace amounts of toxic materials. This problem can be avoided by using recombinantly produced D15 outer membrane protein in heterologous systems which can be isolated from the host in a manner to minimize toxins in the purified material.

35 Particularly desirable hosts for expression in this regard include Gram positive bacteria which do not have

lipopolysaccharide (LPS) and are, therefore, endotoxin free. Such hosts include species of <u>Bacillus</u> and may be particularly useful for the production of non-pyrogenic D15 outer membrane protein, fragments or analogs thereof.

BIOLOGICAL DEPOSITS

Certain plasmids that contain at least a portion coding for a D15 outer membrane protein from strains of Haemophilus influenzae that are described and referred to herein have been deposited with the American Type Culture Collection (ATCC) located at Rockville, Maryland USA 10 pursuant to the Budapest Treaty and prior to the filing of this application. Samples of the deposited plasmids will become available to the public upon grant of a patent based upon this United States patent application. The invention described and claimed herein is not to be 15 limited in scope by plasmids deposited, since the deposited embodiment is intended only as an illustration of the invention. Any equivalent or similar plasmids that encode similar or equivalent antigens as described 20 in this application are within the scope of invention.

DEPOSITE SUMMARY

Clone H. influenzae ATCC Date Designation Deposited 25 DS-712-2-1 Eagan 75604 November 4, 1993 JB-1042-5-1 SB33 75606 November 4, 1993 DS-880-1-2 Eagan 75605 November 4, 1993

The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific Examples. These Examples are described solely for purposes of illustration and are not intended to limit the scope of

20

30

35

the invention. Although specific terms have been employed herein, such terms are intended in a descriptive sense and not for purposes of limitations. Immunological and recombinant DNA methods may not be explicitly described in this disclosure but are well within the scope of those skilled in the art.

EXAMPLES

Methods of molecular genetics, protein biochemistry, and immunology used but not explicitly described in this disclosure and these EXAMPLES are amply reported in the scientific literature and are well within the ability of those skilled in the art.

Example 1

This Example illustrates the cloning and sequencing of the D15 genes.

Genomic DNA was purified from the Haemophilus influenzae type b strain Ca by lysis of the bacteria with pronase and sodium dodecylsulphate followed by phenol extraction and isopropanol precipitation, according to Berns and Thomas, 1965. The DNA was then partially digested with EcoRI and the DNA fraction containing 6-10 kb fragments was isolated following electrophoresis in low-melting point agarose. These fragments were ligated into a lambda gt11 Amp1 vector (Thomas and Rossi, 1986) and cloned as a lysogen into E. coli strain BTA282. Recombinant clones were selected for their ampicillin resistance conferred by the vector. To identify clones producing H. influenzae type b antigen, the clones were replica-plated on nitrocellulose filters and duplicate colonies induced for expression by temperature switch to 42°C for 2 hours. Colonies were lysed by wetting the filters with 1% sodium dodecylsulphate (SDS). The filters were then placed into a chloroform-saturated atmosphere for 15 min. The filters were then assayed by colony radioimmuno-assay using a hyperimmune rabbit anti-H. influenzae type b antiserum absorbed with E. coli lysate

for antigen expression. Clones shown by autoradiography to be producing <u>H. influenzae</u> type b antigens were further purified and their replicates retested for reactivity with the hyperimmune anti-<u>H. influenzae</u> type b antiserum. The antiserum absorbed with 10¹⁰ intact <u>H. influenzae</u> type b bacteria (strain Ca) was used as negative control.

A number of clones were identified which reacted with the unabsorbed, but not with the absorbed antiserum and were further analysed. One of the clones, D15, was 10 purified, grown and found to produce a H. influenzae type b antigen which migrated in sodium dodecyl sulphate polyacrylamide gels with a Mr of about 80 kDa. Lysates from the D15 clone were coupled to Sepharose™ 4B gel and 15 used to affinity-purify anti-D15 antibodies. procedure is described by Thomas et al, 1990, except that the apparent M_r was initially reported to be about 103 kDa. The affinity-purified antibodies to D15 were then shown to react with an $M_{\rm r}$ 80 kDa protein in an outer membrane protein preparation of <u>H. influenzae</u> type b 20 (sarcosyl insoluble fraction - Carlone et al, 1986). Radioimmuno dot blots and Western blots analyses of membrane preparations from both type b and nontypeable Haemophilus influenzae strains showed that affinitypurified anti-D15 antibodies reacted with all isolates. 25 These antibodies were found to be capable of passively protecting infant rats from bacteraemia following intraperitoneal injection of live H. influenzae type b specificity of the protection was The confirmed by absorbing out the protective activity of 30 anti-D15 antibodies with a lysate of E. coli expressing D15 coupled to Sepharose. The protection studies have been described in detail by Thomas et al, 1990.

DNA from the lambda gtll Ampl D15 phage was isolated and a 5.7 kb fragment was released by EcoRI digestion. This fragment was subcloned into pUC19 and the resulting

15

20

25

30

plasmid transformed into <u>E. coli</u> HB101. Recombinant bacteria were found to produce the expected M_r 80 kDa <u>H. influenzae</u> type b antigen when examined by Western blotting. The insert DNA was then characterised by restriction endonuclease mapping. A 2.8 kb <u>HindIII-EcoRI</u> fragment was subcloned into pUC19 to generate plasmid pUC19/D15, which was transformed into <u>E. coli</u> HB101. The recombinant bacteria expressed a M_r 80 kD protein recognized by D15-specific antibodies on Western blot analysis of <u>E. coli</u> lysates.

Plasmid DNA was prepared from two individual colonies of recombinant E. coli HB101 containing the pUC19/D15 plasmid using standard Oligonucleotide sequencing primers of 17-25 bases in length were synthesized on the ABI model 380B Synthesizer and purified by chromatography using OPC cartridges obtained from Applied Biosystems Inc., and used in accordance with the manufactures recommendations. Samples were sequenced using the ABI model 370A DNA Sequencer and dye terminator chemistry according to manufacturers' protocols. This sequence analysis indicated that the D15 gene contains an open reading frame encoding for 789 amino acids, including a putative signal sequence (Figure 1). The derived amino acid sequence was found to contain the sequence of an internal peptide obtained by thrombin digestion of native D15 that had been chemically determined. The amino composition of D15 derived from the D15 gene sequence was comparable (within experimental error) to that of the native protein as determined by amino acid analysis.

Example 2

This Example illustrates the preparation of chromosomal DNA from <u>Haemophilus influenzae</u> strains Eagan, MinnA, SB33, and PAK 12085.

H. influenzae strains were grown on Mueller-Hinton agar or in brain heart infusion broth as described by Harkness et al., 1992.

Eagan chromosomal DNA

Bacteria from 50 mL of culture were pelleted by 5 centrifugation at 5,000 rpm, 20 minutes, 4°C. The pellet was resuspended in 25 mL TE (10mM Tris, 1mM EDTA, pH 8.0) and x 5mL aliquots used for chromosomal preparation. To each aliquot were added 0.6 mL of 10% sarkosyl and 0.15 mL of 20mg/mL proteinase K and the 10 samples incubated at 37°C for 1 hour. The lysate was extracted once with Tris-saturated phenol (pH 8.0) and three times with chloroform:isoamyl alcohol (24:1). aqueous phase was pooled for a final volume of 7 mL. Then, 0.7 mL of 3M sodium acetate (pH 5.2) and 4.3 mL of 15 isopropanol were added to precipitate the DNA which was spooled, rinsed with 70% ethanol, dried, and resuspended in 1 mL of water.

MinnA, SB33, and PAK 12085 chromosomal DNA

Bacteria from 50 mL of culture were pelleted by 20 centrifugation at 5,000 rpm for 15-20 minutes, at 4°C, in Sorvall RC-3B centrifuge. The cell pellet was resuspended in 10 mL of TE (10mM Tris-HCl, 1mM EDTA, pH 7.5), pronase was added to 500 $\mu g/mL$, and SDS to 1%. The sample was incubated at 37°C for about 4 hours until a 25 clear lysate was obtained. The lysate was extracted once with Tris-saturated phenol, once with Tris-saturated phenol/chloroform (1:1), and once with chloroform. final aqueous phase was dialysed for 24 hours against 2 x 500 mL of 1M NaCl at 4°C, changing the buffer once, and 30 for 24 hours against 2 x 500 mL of TE at 4°C, changing the buffer once. The final dialysate was aliquotted for subsequent use.

Example 3

This Example illustrates the preparation of Haemophilus influenzae chromosomal libraries.

H. influenzae Eagan and PAK 12085 chromosomal DNAs were digested with Sau3A I (0.5 unit/10 μg DNA) at 37°C for 15 minutes and size-fractionated by agarose gel electrophoresis. Gel slices corresponding DNA fragments of 15-23 kb were excised and DNA was electroeluted overnight in dialysis tubing containing 3 mL of TAE (40mM Tris-acetate, 1mM EDTA, pH 8.0) at 14V. The DNA was precipitated twice and resuspended in water before overnight ligation with EMBL3 BamH I 10 (Promega). The ligation mixture was packaged using the Lambda in vitro packaging kit (Amersham) according to the manufacturer's instructions and plated onto E. coli NM539 The library was titrated, then amplified and stored at 4°C under 0.3% chloroform.

15 MinnA chromosomal DNA (10 μ g) was digested with Sau3A I (40 units) for 2, 4, and 6 minutes then sizefractionated on a 10-30% sucrose gradient in TNE (20mM Tris-HCl, 5mM NaCl, 1mM EDTA, pH 8.0). containing DNA fragments >5 kb were pooled 20 precipitated. In a second experiment, chromosomal DNA $(2.6 \mu g)$ was digested with Sau3A I (4 units) for 1, 2, and 3 minutes and size- fractionated by preparative agarose gel electrophoresis. Gel slices containing DNA fragments of 10-20 kb were excised and DNA extracted by 25 a standard freeze/thaw technique. The size-fractionated DNA from the two experiments was pooled for ligation with BamH I arms of EMBL3 (Promega). The ligation mix was packaged using the Gigapack II packaging kit (Amersham) and plated on <u>E. coli</u> LE392 cells. The library was titrated, then amplified and stored at 4°C under 0.3% 30 chloroform.

SB33 chromosomal DNA (20 μg) was digested with Sau3A I (40 units) for 2, 4, or 6 minutes and size-fractionated on a 10-30% sucrose gradient in TNE (20mM Tris-HCl, 5mM NaCl, 1mM EDTA, pH 8.0). Fractions containing fragments >5 kb were pooled. In a second experiment, SB33

35

chromosomal DNA (2 μ g) was digested with Sau3A I (4 units) for 2, 4, or 6 minutes and size-fractionated on a preparative agarose gel. Gel slices containing DNA fragments of 10-20 kb were excised and DNA extracted by a standard freeze/thaw technique. The size-fractionated DNA from both experiments was pooled for ligation with BamH I arms of EMBL3 (Promega). The ligation mix was packaged using the Gigapack II packaging kit and plated on LE392 cells. The library was titrated, then amplified and stored at 4°C under 0.3% chloroform.

Example 4

This Example illustrates the screening of the DNA libraries.

The Eagan, MinnA, SB33, and PAK 12085 DNA libraries 15 were plated onto LE392 cells on NZCYM plates using 0.7% top agarose in NZCYM as overlay. Plaque lifts onto nitrocellulose filters were performed following standard procedures, and filters were processed and hybridized with a digoxigenin-labelled D15 probe prepared according 20 the manufacturer's specifications (Boehringer Mannheim). The probe was the <a>EcoR I/<a>Hind III fragment from pUC19/D15 containing the entire Ca D15 gene (Figure Putative plaques were plated and submitted to a second round of screening using the same procedures. 25 Phage DNA was prepared from 500 mL of culture using standard techniques, the insert DNA was excised by Sal I digestion, and cloned into pUC to generate clones DS-712-2-1 (Eagan), DS-691-1-5 (MinnA), JB-1042-5-1 (SB33), and JB-1042-9-4 (PAK 12085), which are shown in Figure 2.

The nucleotide sequences of the D15 genes from <u>H. influenzae</u> type b strains Eagan and MinnA the non-typeable <u>H. influenzae</u> strains SB33 and PAK 12085 were determined and compared with that for strain Ca, as seen in figures 1b, 1C, 1D, 1E and 1F. The desired amino acid sequence are shown in Figures 1B, 1C, 1D and 1E and are

10

15

20

25

compared with the amino acid sequence of the D15 protein of <u>H. influenzae</u> type b Ca (Figure 3).

Example 5

This Example illustrates the expression of rD15 protein in <u>E. coli</u>.

A 2.8 kb fragment <u>HindIII-Eco</u>RI was subcloned into pUC19 and this pUC19/D15 plasmid was transformed into <u>E. coli</u> HB101. Upon induction, the positive clones expressed an 80 kDa protein which was recognized by D15-specific antisera on Western blot analysis. A <u>HindIII-Pst</u> I fragment was also subcloned into pUC19 and shown to express a 67 kDa protein. According to the restriction map, this 67 kDa protein corresponded to a C-terminal truncated D15 protein. On Western blot analysis, this truncated D15 was still recognized by the D15-specific antisera.

Plasmids to express the D15 gene of the non-typeable strain SB33 in <u>E. coli</u> were constructed. Plasmid JB-1042-5-1 containing the SB33 D15 gene and its flanking regions, was digested with <u>Eco</u>R I and <u>Hind</u> III and the 3kb D15 insert subcloned into pUC to give plasmid pRY-60-1 (Figure 4). Appropriate oligonucleotides were synthesized to restore the native D15 sequence between the ATG codon of the expression plasmid pT7-7 and the BsrF I site within the D15 gene. These oligonucleotides had the following sequence:

Nde

5'- TATGGCACCTTTTGTGGCAAAAGATATTCGTGTGGATGGTGTTCAAGGTG

ACCGTGGAAAACACCGTTTTCTATAAGCACACCTACCACAAGTTCCACTGAATCT

ACTTAGAATCAACAAACCGAGCAAGTTTACCTGTTCGTG - SEQ ID NO: 50

TGGTTGTTTAGGCTCGTTCAAATGGACAAGCACGGCC-5'- SEQ ID NO: 51

Berf I

Plasmid pRY-60-1 was digested with <u>Eco</u>R I and <u>Bsr</u>F I and the DNA fragment containing most of the D15 gene was purified. pUC was digested with <u>Eco</u>R I and <u>Nde</u> I and the

15

vector fragment purified. A multi-component ligation between the pUC and D15 fragments and the oligonucleotides generated plasmid DS-860-1-1 which contains a D15 sequence without a promoter. pT7-7 was digested with Nde I and EcoR I and the vector fragment purified. DS-860-1-1 was digested with Nde I and EcoR I and the D15 insert was purified and ligated with the T7-7 vector generating plasmid DS-880-1-2 (Figure 4).

The plasmid constructions were performed using <u>E. coli</u> JM109 as host. For expression, plasmid DS-880-1-2 was transformed into <u>E. coli</u> BL21/DE3, BL21/DE3/pLysS, or JM109/DE3 cells. Transformation of the cells was performed using either calcium chloride-treated competent cells or by electroporation using a BioRad electroporator. Transformed cells were grown in YT, M9, or NZCYM media and induced with IPTG or other inducing agents.

Example 6

This Example illustrates the construction and 20 expression of the GST-D15 fragment hybrid gene in \underline{E} . coli.

A forward sense primer (primer 1) 5'-GGGGAATTCCAAAAGATGTTCGT (SEQ ID NO: 52) and a reverse antisense primer CACGAATTCCCTGCAAATC-5' (primer 7 - SEQ 25 ID NO: 53) were used to amplify a 2.8 Kb fragment <u>Hind</u>III-<u>Eco</u>RI of the D15 gene by the polymerase chain reaction that encodes the N-terminal amino acid residues 22 to 223 of the primary sequence of D15 protein (Figure The nucleotide sequence of the 609bp amplified fragment was confirmed by DNA sequencing. The amplified 30 segment was ligated into the pGEX-2T vector downstream from the GST gene and transformed into E. coli Colonies expressing the H. influenzae type b antigen were screened with a rabbit anti-H. influenzae type b antiserum by colony radioimmunoassay and isolated. 35 The glutathione-S-transferase-D15 fragment fusion protein

15

25

30

35

produced by transformed \underline{E} , \underline{coli} was isolated by affinity purification on glutathione agarose.

Example 7

This Example describes alternative expression 5 systems for rD15.

The D15 gene or fragments thereof are also expressed E. coli under the control of other regulated promoters. The D15 gene or fragments thereof are expressed in the absence of the leader peptide, or in other cloning systems where toxicity of D15 expression to the host is not problematic. The gene or fragments thereof are synthesized de novo or by employing the polymerase chain reaction using suitable primers. genes are cloned into suitable cloning vectors bacteriophage vectors in E. coli or other suitable hosts directly when toxicity can be avoided. Expression systems are Gram-positive bacteria (such as Bacillus species), pox virus, adenovirus, baculovirus, yeast, fungi, BCG or mammalian expression systems.

20 Example 8

This Example illustrates the protocol for extraction and purification of rD15 from $\underline{E.\ coli}$ expression system.

The cell pellet from a 250 mL culture, prepared as described in Example 5, was resuspended in 40 mL of 50 mM Tris, pH 8.0, and disrupted by sonication (3 x 10 min, 70% duty circle). The extract was centrifuged at 20,000 x g and the resulting pellet saved. The initial pellet was re-extracted with 40 mL of 50 mM Tris, 0.5% Triton X-100, 10 mM EDTA, pH 8.0. The suspension was then sonicated for 10 minutes at 70% duty circle. The extract was centrifuged at 300 x g for 5 minutes. The resulting supernatant was centrifuged again at 20,000 x g for 30 min and the resulting pellet was saved. The pellet was resuspended in 50 mM Tris, 0.5% Triton X-100, 10 mM EDTA, pH 8.0. The suspension was then mixed with PBS/ 8 M urea to a final urea concentration of 6 M. The solution was

20

25

30

35

then dialyzed against PBS to remove urea. After dialysis, the solution was centrifuged at 300 x g for 10 min., the supernatant was saved and stored at 4°C.

Example 9

This Example demonstrates the purification of GST-(D15 fragment) fusion protein using glutathione-Sepharose 4B affinity chromatography.

Five mg of GST-(D15 fragment) fusion protein crude extract, prepared as described in Example 6, were dissolved in 5 mL of phosphate buffer saline (PBS) containing 1% Triton X-100. The solution was then loaded Glutathione-Sepharose 4B column mL) equilibrated with PBS containing 1% Triton X-100. The run-through of the column was discarded. The column was washed with 20 mL of PBS and the GST-(D15 fragment) fusion protein was eluted with 50 mM Tris-HCl buffer, pH 8.0, containing 5 mM glutathione. Elution was monitored by absorbance at 280 nm. Protein-containing fractions (2 mL/fraction) were collected and pooled. The purity of the protein was assessed by SDS-PAGE (Figure 9, lane 3). The final volume of the purified fusion protein was 6 mL. Example 10

This Example illustrates the protocol used for thrombin digestion of proteins to release the truncated D15 molecule.

The GST-(D15 fragment) fusion protein sample from Example 9 (0.1 to 0.5 mg protein/mL) was dialyzed against 1 L of 50 mM Tris-HCl buffer (pH 8.5) 3 times with at least 2 hour intervals at 4°C to remove protease inhibitors. After dialysis, the solution was treated with human thrombin (Sigma) at a ratio of 1 mL of solution to 25 units of thrombin. The cleavage reaction was carried out at 37°C for 2 hr and analysed by SDS-PAGE (Figure 9, lane 4). The reaction was stopped by placing the solution in ice.

Example 11

15

20

25

30

This Example illustrates the procedure used for N-terminal rD15 fragment purification from GST using Glutathione-Sepharose 4B affinity chromatography.

A thrombin-digested GST-(D15 fragment) sample, prepared as described in Example 10, was loaded onto a Glutathione-Sepharose 4B column (2 mL) equilibrated with PBS containing 1% Triton X-100. The run-through of the column containing the N-terminal rD15 fragment was saved. After washing the column with 20 mL of PBS, the affinity column was regenerated by removing GST using 50 mM Tris-HCl buffer, pH 8.0, containing 5 mM glutathione. The purity of rD15 fragment was analysed by SDS-PAGE (Figure 9, lane 5). This N-terminal rD15 fragment contains amino acids 63-223 of the D15 protein as a result of cleavage at the spacious thrombin site shown in Figure 1A.

Example 12

This Example illustrates the protocol used for the purification of D15-specific polyclonal antibodies by affinity chromatography using GST-(D15 fragment) fusion protein.

The recombinant GST-(D15 fragment) fusion protein, prepared as described in Example 9, was conjugated to cyanogen bromide-activated Sepharose. The affinity column was then used to purify antibodies from a rabbit hyperimmune anti-H. influenzae type b antiserum. The affinity purified-antibodies were shown by immunoblotting to react with a 80 kDa component present in the lysates of E. coli transformed with pUC9/D15 and in the lysates of several typeable and nontypeable H. influenzae isolates. These results confirmed that the DNA segment encoding the D15 fragment of the fusion protein was part of the open reading frame of the D15 gene.

Similarly, antisera raised against the recombinant fusion protein (Example 9) or the purified N-terminal rD15 fragment (Example 11) reacted with the D15 protein produced by H. influenzae strains (Example 13).

10

15

20

25

35

Example 13

This Example describes the protocol used for the purification of native D15 from <u>H. influenzae</u>.

Cell paste of the non-typeable <u>H. influenzae</u> SB33 strain, prepared from a culture grown in brain heart infusion medium supplemented with NAD $(2\mu g/mL)$ and HEMIN $(2\mu g/mL)$ at 37°C, as described in Panezutti, et al, 1993, was resuspended in 50 mM Tris-HCl, pH 8.0, containing 0.5% Triton X-100 and 10 mM EDTA (20 mL per 1 g of cell paste). The mixture was stirred at room temperature for 2 hr, then centrifuged at 20,000 x g for 30 minutes. The D15 was located in the supernatant and further purified.

Purification of native D15 was achieved by affinity chromatography using a D15-specific monoclonal antibody (see Example 24). The D15 extract (25 mL) was mixed with the affinity matrix (1 mL) at room temperature for 2 hr. The mixture was packed into a column and the run-through was discarded. The column was sequentially with the following buffers: 50 mM Tris-HCl, pH 8.0, containing 0.5% Triton X-100 and 10 mM EDTA; 1 M HEPES buffer, pH 6.8; 50 mM Tris-HCl, pH 8.0, containing 0.5% Triton X-100 and 10 mM EDTA; and 10 mM phosphate buffer, pH 8.0. D15 was then eluted from the column with 3 mL of 50 mM diethylamine, pH 12.0 and the protein solution was neutralized by 1 M HEPES, pH 6.8 (1/10 volume). The affinity-purified native D15 was analysed by SDS-PAGE and stored at -20°C.

Example 14

This Example describes the procedure used for the preparation of D15-PRP conjugates.

Haemophilus influenzae type b oligosaccharides (PRP) prepared by controlled acid hydrolysis were conjugated either with the purified native (Example 13) or recombinant D15 (Example 8) as well as with its fragments (Example 11) using periodate oxidation as described in US Patent 4356170 and further details of which are presented

10

15

20

25

30

35

in Example 17. The mean molecular size of the PRP molecules used for conjugation was determined as being approximately 20,000 Daltons. The conjugation was carried out without a linker molecule but may also be carried out with a linker molecule. A PRP/D15 molar ratio of approximately 7 was used to provide an excess of PRP hapten.

The PRP/rD15 conjugate was tested according to the protocol of Example 18 for immunogenicity in rabbits and elicited both primary and secondary anti-PRP IgG and anti-D15 antibody responses (Table 9). Rabbit anti-rD15-PRP antisera also strongly reacted with both native D15 and rD15 as judged by immunoblot analysis. These data indicate that rD15 can be used as a carrier protein in a conjugate vaccine. In addition, a rD15-PRP conjugate vaccine should ensure a more consistent protection against H. influenzae type b disease, particularly in infants, as a result of the additional homotypic protection provided by antibodies directed against the D15 protein.

Example 15

This Example describes the preparation of D15 peptides.

D15 peptides (Table 2) were synthesized using an ABI 430A peptide synthesizer and optimized t-Boc chemistry as described by the manufacturer, then cleaved from the resin by hydrofluoric acid (HF). The peptides were purified by reversed-phase high performance liquid chromatography (RP-HPLC) on a Vydac C4 semi-preparative column (1 x 30 cm) using a 15 to 55% acetonitrile gradient in 0.1% trifluoryl acetic acid (TFA) developed over 40 minutes at a flow rate of 2 mL/min. All synthetic peptides (Table 2) used in biochemical and immunological studies were >95% pure as judged by analytical HPLC. Amino acid composition analyses of

these peptides performed on a Waters Pico-Tag system were in good agreement with their theoretical compositions.

Example 16

This Example describes the protocol used for D15 peptide-specific antisera production.

Guinea pigs and rabbits were immunized individual peptides (50 to 200 μ g) emulsified with Freund's complete adjuvant and injected intramuscularly. After two booster doses with the same amount of peptide in incomplete Freund's adjuvant at +14 and +28 days, the anti-peptide antisera were collected on day +42 and tested by ELISAs and immunoblotting. Both rabbit and guinea pig antisera were shown to be monospecific for their respective immunizing peptides by the peptidespecific ELISAs (Table 6). In addition, both guinea pig and rabbit antisera raised against D15 peptides reacted with both <u>H. influenzae</u> type b and non-typeable D15 on immunoblot analyses. Since most D15 peptides induced strong anti-peptide antibody responses in at least one animal species, they are appropriate immunogens to be included in immunogenic compositions including vaccine preparations.

Example 17

5

10

15

20

30

35

This Example describes the procedure used for the preparation of PRP-BSA conjugates.

0.5 mL of periodate-oxidized PRP (25mg in 1 mL of 0.1 M sodium phosphate buffer, pH 6.0), prepared from native PRP treated with aqueous periodic acid (Carlone et al, 1986), was added to bovine serum albumin (BSA) (1.32 mg; 0.02 μ mol) in 0.5 mL of 0.2 M sodium phosphate buffer, pH 8.0, followed by the addition of sodium cyanoborohydride (14 μ g; 0.22 μ mol; 10 eqv. to BSA). After incubation at 37°C for 5 days, the reaction mixture was dialysed against 4 L of 0.1 M phosphate buffer, pH 7.5. The resulting solution was applied onto an analytical Superose 12 column (15 x 300 mm, Pharmacia)

equilibrated with 0.2 M sodium phosphate buffer, pH 7.2, and eluted with the same buffer. Fractions were monitored for absorbance at 230 nm. The first major protein peak was pooled and concentrated in a Centriprep 30 to 2.2 mL.
5 The amount of protein was determined using the Bio Rad protein assay, and was found to be 300 μg/mL. The presence of PRP in the protein conjugate fraction was confirmed by the Orcinol test.

Example 18

This Example describes the protocol used for the production of anti-PRP antisera in animals using rD15-PRP conjugates.

Rabbits were immunized intramuscularly with rD15-PRP conjugates (Example 14) (5 to 50 μ g PRP equivalent) mixed with 3 mg AlPO₄ per mL, followed by two booster doses (half amount of the same immunogen) at 2 week intervals. Antisera were collected every 2 weeks after the first injection, heat-inactivated at 56°C for 30 minutes and stored at -20°C.

20 Example 19

15

25

30

35

This Example illustrates the reactivity between D15 peptides and anti-peptide and D15-specific antisera using D15-specific and peptide-specific ELISAs.

Microtiter wells (Nunc-Immunoplate, Nunc, Denmark) were coated with 200 ng of purified rD15 or 500 ng of individual peptides in 50 μ L of coating buffer (15 mM Na₂CO₃, 35 mM NaHCO₃, pH 9.6) for 16 hours at room temperature. The plates were then blocked with 0.1% (w/v) BSA in phosphate buffer saline (PBS) for 30 minutes at room temperature. Serially diluted antisera were added to the wells and incubated for 1 hour at room temperature. After removal of the antisera, the plates were washed five times with PBS containing 0.1% (w/v) Tween-20 and 0.1% (w/v) BSA. F(ab')₂ fragments from goat anti-rabbit, guinea pig, mouse, or human IgG antibodies conjugated to horseradish peroxidase (Jackson ImmunoResearch Labs Inc.,

15

PA) were diluted (1/8,000) with washing buffer, and added onto the microtiter plates. After 1 hr incubation at room temperature, the plates were washed five times with the washing buffer. The plates were then developed using the substrate tetramethylbenzidine (TMB) in H₂O₂ (ADI, Toronto). The reaction was stopped with 1N H2SO4 and the optical density was measured at 450 nm using a Titretek Multiskan II (Flow Labs., Virginia). Two irrelevant peptides as negative controls in the peptide-specific ELISAs. Assays were performed in triplicate, and the reactive titer of each antiserum was defined as the dilution consistently showing 2-fold increase absorbance value over those obtained from the negative controls. The results obtained are summarized in Tables 3, 6 and 8 and in the DETAILED DESCRIPTION OF THE INVENTION above.

Example 20

This Example illustrates the measurement of the anti-PRP IgG titers in rabbit anti-PRP-D15 conjugate antisera using a PRP-specific ELISA.

20 Microtiter wells (Nunc-Immunoplate, Nunc, Denmark) were coated with 200 ng of purified PRP-BSA (see Example 17) in 200 μ L of coating buffer (15 mM Na₂CO₃, 35 mM NaHCO₃, pH 9.6) for 16 hours at room temperature. plates were then blocked with 0.1% (w/v) BSA in phosphate 25 buffer saline (PBS) for 30 minutes at room temperature. Serially diluted rabbit antisera raised against PRP-D15 conjugates were added to the wells and incubated for 1 hour at room temperature. After removal of the antisera, the plates were washed five times with PBS containing 30 0.1% (w/v) Tween-20 and 0.1% (w/v) BSA. $F(ab')_2$ fragment from goat anti-rabbit IgG antibodies conjugated to horseradish peroxidase (Jackson ImmunoResearch Labs Inc., PA) were diluted (1/8,000) with washing buffer, and added onto the microtiter plates. After 1 hour incubation at room temperature, the plates were washed five times with 35 the washing buffer. The plates were then developed using

10

15

20

25

30

35

the substrate tetramethylbenzidine (TMB) in H_2O_2 (ADI, Toronto). The reaction was stopped with 1N H_2SO_4 and the optical density measured at 450 nm using a Titretek Multiskan II (Flow Labs., Virginia). A standard anti-PRP antiserum of known titer was included as positive control. Assays were performed in triplicate, and the reactive titer of each antiserum was defined as the reciprocal of the dilution consistently showing a 2-fold increase in O.D. value over that obtained with the pre-immune serum (Table 9).

Example 21

This Example describes the protocol used for the production of D15-specific antisera using purified D15, rD15 or N-terminal rD15 fragment.

New Zealand White rabbits (Maple Lane) and guinea pigs (Charles River) were immunized intramuscularly (IM) with a 10 μ g dose of either affinity-purified native D15 (Example 13), recombinant D15 (Example 8) or N-terminal rD15 fragment (Example 11) emulsified in Freund's complete adjuvant (Difco). Animals were boosted on day 28 with another 10 μ g dose of affinity-purified D15 or rD15 or rD15 fragment emulsified in Freund's incomplete adjuvant and bled on day 42 via the marginal ear vein. D15-specific polyclonal antibodies were purified from this material as described in Example 12.

Example 22

This Example illustrates the protective activity of D15-specific antisera against <u>H. influenzae</u> type b challenge using the infant rat model of bacteremia.

Five-day old infant rats were inoculated subcutaneously (SC) on the dorsum with 0.15 mL of two different rabbit anti-N-terminal rD15 fragments. Preimmune sera were used as negative controls. One day after immunization, the infant injected rats were intraperitoneally (IP) with 200 colony-forming units (cfu) of <u>Haemophilus influenzae</u> type b Minn A strain (0.1

ml) freshly grown in brain heart infusion (BHI) medium supplemented with cofactors and diluted in PBS containing 0.5 mM MgCl₂ and 0.15 mM CaCl₂. One day later, blood samples were collected via cardiac puncture under methoxyflurane anaesthesia and plated on chocolate agar plates. The number of bacteria per mL of blood was quantified after 24 hr. The statistical significance of differences observed in the levels of bacteremia relative to controls was analyzed by the Student's t-test. The results are summarized in Table 1.

Example 23

5

10

This Example describes the protocol used for the generation of D15-specific T-cell lines.

BALB/c (H-2d) mice purchased from Charles River Animal Farm (Montreal, Canada) were individually primed 15 subcutaneously with 20 μg of rD15 adsorbed to 1.5 mg of aluminium phosphate (alum). The animals were boosted twice with the same dose of immunogen at 3 week intervals. Ten days after the final boost, spleens of 20 immunized mice were removed. Splenocytes were cultured at 5.75 x 10^5 cells per well in a final volume of 200 μL of RPMI 1640 medium (Flow Lab.) supplemented with 10% heatinactivated fetal calf serum (Gibco), 2 mM L-glutamine (Flow Lab.), 100 U/mL) penicillin (Flow Lab.) and 5 \times 10⁻⁵ 25 M 2-mercaptoethanol (Sigma) in the presence of varying concentrations (1, 10 and 100 μ g per mL) of individual D15 peptides (Table 2) in 96-well plates (Nunc, Denmark). Cultures were kept in a humidified incubator in the presence of 5% CO₂/air. Triplicate cultures were 30 performed for each concentration of each peptide. days later, 150 μ L of 10% rat concanavalin A culture supernatant diluted in culture medium was added to the microtiter plate wells as a source of Interleukin-2 (IL-2) to expand peptide-specific T-cells. Six days later, 35 μ L of supernatant were removed microculture, and 150 μ L of fresh IL-2 containing culture

10

15

20

25

30

35

supernatant added to further expand and maintain the viability of the peptide-specific T-cells. After a further 6 day-incubation, the cells were washed three times, each time with 200 μL of culture medium.

Each set of cultures was then stimulated with the corresponding concentrations (1, 10 and 100 μ g per mL) of the peptide in the presence of 2×10^5 irradiated (1,500 rad) BALB/c spleen cells in a final volume of 200 μL of culture medium. Sixty μL of supernatant were then removed from each microculture. The supernatants from each triplicate cultures set were pooled. All supernatants were assayed for IL-2, Interleukin-4 and Interferon-gamma Detections of IL-2 and IL-4 were performed using murine IL-2 and IL-4 ELISA kits purchased from Endogen Inc. (MA, USA) respectively. Assay of IFN- γ was performed using a mouse IFN- γ ELISA kit supplied by Genzyme Corporation (MA, USA). Test culture supernatants were assayed at 1 in 5 dilution according to the manufacturers' instructions. The results obtained are set forth in Table 7.

Example 25

This Example describes the general procedure used for the production of murine D15-specific monoclonal antibodies.

BALB/c mice were immunized intraperitoneally with 20 to 50 μ g of the N-terminal rD15 fragment (Example 11) emulsified in Freund's complete adjuvant. Two weeks later, the mice were given another injection of the same amount of immunogen in incomplete Freund's adjuvant (IFA). Three days before the fusion, the mice were boosted again with the same amount of immunogen in IFA. Hybridomas were produced by fusion of splenic lymphocytes from immunized mice with non-secreting Sp2/0 myeloma cells as previously described by Hamel et al. (1987). D15-specific hybridomas were cloned by sequential limiting dilutions and screened for anti-D15 monoclonal

antibody production. Eight D15-specific hybridoma cell lines were identified, expanded and frozen in liquid nitrogen. One of the hybridoma cell lines, 6C8-F6-C6, has been partially characterized. The monoclonal antibody (MAb 6C8-F6-C6) reacts with peptide D15-P8. This MAb 6C8-F6-C6 was used to prepare the D15-specific MAb affinity column and purify native D15 from H. influenzae cell paste (Example 13).

57

TABLE 1

PROTECTIVE EFFECT OF PASSIVELY TRANSFERRED ANTI-N-TERMINAL RD15
FRAGMENT ANTIBODIES IN THE INFANT RAT MODEL OF BACTEREMIA¹

	cfu/0.1 mL blood		
Rabbit antisera	Pre-immune	Post-immunization	p value
Rb#434	510 (6/6)2	6 (1/6)	<0.001
Rb#435	910 (4/4)	6 (1/4)	<0.001

Five-day old infant rats were passively immunized with 0.15 mL of rabbit anti-N-terminal rD15 fragment s.c. One day later, the infant rats were challenged with 200 cfu of <u>H. influenzae</u> type b strain MinnA (0.1 mL, IP). The blood samples were taken from each rat 24 hours after the challenge and analysed for bacteria counts.

The parentheses indicate the number of rats found to be bacteremic out of the total number of rats challenged.

58

TABLE 2

SEQUENCE OF OVERLAPPING SYNTHETIC PEPTIDES ENCOMPASSING
THE ENTIRE D15 ANTIGEN SEQUENCE

PEPTIDES	RESIDU	ES SEQUENCES	SEQ ID NO:
D15-P1	20-49	APFVAKDIRVDGVQGDLEQQIRASLPVRAG	14
D15-P2	45-74	PVRAGQRVTDNDVAMIVRSLFVSGRFDDVK	15
D15-P3	68-99	GRFDDVKAHQEGDVLVVSVVAKSIISDVKIKG	16
D15-P4	93-122	SDVKIKGNSVIPTEALKQNLDANGFKVGDV	17
D15-P5	114-143	angfkvgdvlireklnefaksvkehyasvg	18
D15-P6	135-164	VKEHYASVGRYNATVEPIVNTLPNNRAEIL	19
D15-P7	157-187	PNNRAEILIQINEDDKAKLASLTFKGNESVS	20
015-P8	180-209	FKGNESVSSSTLQEQMELQPDSWWKKLWGNK	21
D15-P9	199-228	PDSWWKLWGNKFEGAQFEKDLQSIRDYYLN	22
D15-P10	219-249	LQSIRDYYLNNGYAKAQITKTDVQLNDEKTK	23
D15-P11	241-270	VQLNDEKTKVNVTIDVNEGLQYDLRSARII	24
D15-P12	262-291	YDLRSARIIGNLGGMSAELEPLLSALHLND	25
15-P13	282-312	PLLSALHLNDTFRRSDIADVENAIKAKLGER	26
15-P14	304-333	AIKAKLGERGYGSATVNSVPDFDDANKTLA	27
15-P15	325-354	FDDANKTLAITLVVDAGRRLTVRQLRFEGN	28
15-P16	346-375	VRQLRFEGNTVSADSTLRQEMRQQEGTWYN	29
15-P17	367-396	RQQEGTWYNSQLVELGKIRLDRTGFFETVE	30
15-P18	390-416	GFFETVENRIDPINGSNDEVDVVYKVK	31
15-P19	410-435	DVVYKVKERNTGSINFGIGYGTESGI	32
15-P20	430-455	GTESGISYQASVKQDNFLGTGAAVSI	33
15-P21	450-477	GAAVSIAGTKNDYGTSVNLGYTEPYFTK	34
15-P22	471-497	TEPYFTKDGVSLGGNVFFENYDNSKSD	35
15-P23	491-516	YDNSKSDTSSNYKRTTYGSNVTLGFP	36
15-P24	511-538	VTLGFPVNENNSYYVGLGHTYNKISNF	37
15-P25	532-559	YNKISNFALEYNRNLYIQSMKFKGNGIK	38
15-P26 !	554-582	KGNGIKTNDFDFSFGWNYNSLNRGYFPTK	39
15-P27 !	577-602	GYFPTKGVKASLGGRVTIPGSDNKYYK	40
L5-P28 !	596-625	SDNKYYKLSADVQGFYPLDRDHLWVVSAK	41

D15-P29	619-646	LWVVSAKASAGYANGFGNKRLPFYQTYT	42
D15-P30	641-666	FYQTYTAGGIGSLRGFAYGSIGPNAI	43
D15-P31	662-688	GPNAIYAEYGNGSGTGTFKKISSDVIG	44
D15-P32	681-709	KISSDVIGGNAIATASAELIVPTPFVSDK	45
D15-P33	705-731	FVSDKSQNTVRTSLFVDAASVWNTKWK	46
D15-P34	725-750	VWNTKWKSDKNGLESDVLKRLPDYGK	47
D15-P35	745-771	LPDYGKSSRIRASTGVGFQWQSPIGPL	48
D15-P36	769-798	GPLVFSYAKPIKKYENDDVEOFOFSIGGSF	49

TABLE 3

REACTIVITY OF RABBIT AND GUINEA PIG ANTI-N-TERMINAL rD15

FRAGMENT ANTISERA WITH D15 SYNTHETIC PEPTIDES

		React	ive Titers		
	Rabbit	antisera	Guine	a pig ant	isera
Peptides	3434	435	858	859	860
D15-P1	400	1,600	6,400	6,400	6,400
D15-P2	1,600	<100	100	100	<100
D15-P3	400	<100	100	<100	<100
D15-P4	25,600	6,400	<100	<100	<100
D15-P5	6,400	400	1,600	25,600	400
D15-P6	1,600	6,400	400	6,400	6,400
D15-P7	6,400	1,600	25,600	25,600	25,600
D15-P8	6,400	6,400	25,600	409,600	409,600
D15-P9	<100	<100	400	1,600	1,600
D15-P10	<100	<100	400	6,400	<100

61

TABLE 4

INHIBITION OF ANTI-N-TERMINAL rD15 FRAGMENT ANTIBODY-INDUCED PROTECTION BY D15 PEPTIDES IN THE INFANT RAT MODEL OF BACTEREMIA

Group #	Antibody	cfu / 10 μL blood	cfu in each group/cfu in group #4 (control) (%)
1	Anti-D15 Ab + PBS	60 ± 120 (3/7)	3
2	Anti-D15 Ab + peptides	300 ± 240 (6/7)	13
3	Anti-D15 Ab + rD15	1,520 ± 1,280 (7/7)	64
4	PBS + peptides	2,360 ± 1,200 (6/7)	100

One half mL of rabbit anti-N-terminal rD15 fragment antiserum (Anti-rD15 fragment Ab) was mixed with either nine D15 peptides (100 $\mu \rm g$ of peptides D15-P2 to D15-P10, See TABLE 2) or with 600 $\mu \rm g$ of N-terminal rD15 fragment at room temperature for 1 hr. Antiserum and peptides mixed with PBS were used as controls. Seven-day old infant rats were injected s.c. with 0.2 mL of the various preparations. After 24 h, the infant rats were challenged I.P. with 200 cfu of H. influenzae type b strain MinnA. The blood samples were taken at 24 h after the challenge. The numbers in parentheses indicate the number of animals that were bacteremic out of the total number of animals challenged. The level of bacteremia is expressed as the mean of values obtained from seven infant rats tested individually \pm one standard deviation (SD).

62 TABLE 5.

INHIBITION OF THE IMMUNOPROTECTION ABILITY OF THE RABBIT ANTI-N-TERMINAL rD15 FRAGMENT ANTISERUM ABSORBED WITH D15 PEPTIDES (D15-P4 TO D15-P8) IN THE INFANT RAT MODEL OF BACTEREMIA

Group #	Antibody	cfu / 10 μL blood	cfu in each group/ cfu in group #3 (%)
1	rD15 Ab + PBS	220 ± 360 (3/6)	8
2	rD15 Ab + peptides	2,960 ± 560 (6/6)	106.
3	PBS + peptides	2,800 ± 360 (6/6)	100

One half mL of rabbit anti-rD15 fragment antiserum (rD15 Ab) was mixed with five D15 peptides (peptides P4 to P8, 250 μ g of each peptide) at room temperature for 1 hr. Antiserum and peptides diluted in PBS were used as controls. Seven-day old infant rats were injected s.c. with 0.2 mL of the indicated material. After 24 h, the infant rats were challenged I.P. with 200 cfu of H. influenzae type b strain MinnA. The blood samples were collected 24 h after challenge. The numbers in parentheses indicate the number of animals that were bacteremia out of the total number of animals challenged. The level of bacteremia is expressed as the mean of values obtained from six infant rats tested individually \pm one SD.

63

TABLE 6

REACTIVITY OF RABBIT, GUINEA PIG AND MOUSE ANTI-rD15 ANTISERA WITH D15 PEPTIDES

Peptide	Rabbit ²	Reactive Titer' Guinea Pig'	Mouse ⁴
D15-P1	-	-	+
D15-P2	-	+++	+
D15-P3	-	•	+
D15-P4	+	+	+
D15-P5	-	-	+
D15-P6	-	+	+
D15-P7	-	<u>.</u>	+
D15-P8	-	++++	++++
D15-P9	-	-	+
D15-P10	•		+++
015-P11	-	-	+++
015-P12	-	-	+
015-P13	-	-	+
)15-P14	+++	+	+
15-P15	-	-	+
15-P16	-	-	+
15-P17	-	-	+
15-P18	-	-	+
15-P19	·	-	+
15-P20	-	-	+
15-P21	-	-	+
15-P22	-	-	+
15- P 23		-	+
15-P24	-	-	+
15- P 25	-	-	+
15- P2 6	-	-	+++
L5- P 27	-	, +	+

TABLE 6 (continued)

D15-P28	-	•	, +
D15-P29	-	-	+
D15-P30	-		+
D15-P31	-	•	+
D15-P32	-	-	+
D15-P33	-	-	+
D15-P34	-	-	+
D15-P35	-	-	+
D15-P36	++++	-	+

The reactive titer is based on peptide-specific ELISAs. +, ++, +++, and ++++ represent reactive titers of animal antisera tested at 1/300, 1/1000, 1/2000, and 1/5000 dilutions, respectively; - means nonreactive.

Titer represents the average value of two rabbit antisera raised against rD15.

Titer represents the average value of two guinea pig antisera raised against rD15.

Titer represents the average value of five mouse antisera raised against r15.

TABLE 7
T-CELL STIMULATORY ACTIVITY OF D15 PEPTIDES

Peptide	IL-2 ²	CYTOKINE RELEASE $(pg/mL)^{1}$ γ -IFN ³	IL-4 ⁴
D15-P1	-	-	-
D15-P2	122	-	_
D15-P3	25	-	_
D15-P4	-	-	_
D15-P5	742	38,000	13
D15-P6	-	-	_
D15-P7	-	<u>-</u>	_
D15-P8	-	-	_
D15-P9	-	-	_
D15-P10	108	1,900	_
D15-P11	-	-	-
D15-P12	1,052	6,100	_
D15-P13	105	6,200	- 56
D15-P14	_	-	_
D15-P15	-	-	_
D15-P16	48	_	_
D15-P17	-	-	_
D15-P18	32	4,800	_
D15-P19	882	24,500	_
D15-P20	•	-	_
D15-P21	•	-	_
D15-P22	-	-	
D15-P23	78	-	_
D15-P24	103	-	_
D15-P25	-	-	_
D15-P26	572	6,700	_
D15-P27	274	7,505	68

TABLE 7 (continued)

D15-P28	142	742	-
D15-P29	-	-	-
D15-P30	-	-	_
D15-P31	-	•	-
D15-P32	-	•	_
D15-P33	-	-	-
D15-P34	82	603	-
D15-P35	107	751	-
D15~P36	-	-	-

Results are expressed as mean values of triplicate cultures. All standard deviations were less than 15%. Immunodominant Thi-cell epitopes are highlighted with bold and Thi-cell epitopes are in italics.

TABLE 8

RABBIT AND GUINEA PIG ANTIBODY RESPONSES TO D15 PEPTIDES

Reactive Tite Immunogen D15-P1 D15-P2 D15-P2 D15-P3 D15-P4 D15-P5 D15-P5 D15-P6 D15-P7 D15-P7 D15-P8 D15-P9 D15-P9 D15-P1 D15-P10 D15-P10 D15-P12 D15-P13 NT4	
D15-P1 102,400 D15-P2 204,800 D15-P3 51,200 D15-P4 204,800 D15-P5 51,200 D15-P6 51,200 D15-P7 204,800 D15-P7 204,800 D15-P9 102,400 D15-P10 102,400 D15-P11 51,200 D15-P12 102,400 D15-P13 NT4	er¹
D15-P2 204,800 D15-P3 51,200 D15-P4 204,800 D15-P5 51,200 D15-P6 51,200 D15-P7 204,800 D15-P8 51,200 D15-P9 102,400 D15-P10 102,400 D15-P11 51,200 D15-P12 102,400 D15-P13 NT	Guinea Pig³
D15-P3	819,200
D15-P4 204,800 D15-P5 51,200 D15-P6 51,200 D15-P7 204,800 D15-P8 51,200 D15-P9 102,400 D15-P10 102,400 D15-P11 51,200 D15-P12 102,400 D15-P13 NT	1,637,400
D15-P5 51,200 D15-P6 51,200 D15-P7 204,800 D15-P8 51,200 D15-P9 102,400 D15-P10 102,400 D15-P11 51,200 D15-P12 102,400 D15-P13 NT ⁴	1,637,400
D15-P6 51,200 D15-P7 204,800 D15-P8 51,200 D15-P9 102,400 D15-P10 102,400 D15-P11 51,200 D15-P12 102,400 D15-P13 NT ⁴	819,200
D15-P7 204,800 D15-P8 51,200 D15-P9 102,400 D15-P10 102,400 D15-P11 51,200 D15-P12 102,400 D15-P13 NT ⁴	1,637,400
D15-P8 51,200 D15-P9 102,400 D15-P10 102,400 D15-P11 51,200 D15-P12 102,400 D15-P13 NT	409,600
D15-P9 102,400 D15-P10 102,400 D15-P11 51,200 D15-P12 102,400 D15-P13 NT	819,200
D15-P10 102,400 D15-P11 51,200 D15-P12 102,400 D15-P13 NT	409,600
D15-P11 51,200 D15-P12 102,400 D15-P13 NT ⁴	409,600
D15-P12 102,400 D15-P13 NT ⁴	819,200
D15-P13 NT*	819,200
	204,800
D15 D14	204,800
D15-P14 NT	409,600
D15-P15 NT	204,800
D15-P16 NT	819,200
015-P17 NT	204,800
D15-P18 NT	312,500
D15-P19 NT	312,500
D15-P20 NT	62,500
015-P21 NT	62,500
015-P22 NT	12,500
215-P23 NT	1,562,500
215-P24 NT	312,500
15-P25 NT	62,500

TABLE 8 (continued)

D15-P26	NT	500
D15-P27	NT	1,500
D15-P28	NT	1,250
D15-P29	NT	<500
D15-P30	NT .	<500
D15-P31	NT	<500
D15-P32	NT	12,500
D15-P33	NT	12,500
D15-P34	NT	62,500
D15-P35	NT	1,250
D15-P36	NT	12,500

The reactive titer is based on peptide-specific ELISAs. A titer below 500 indicates that the peptide is not immunogenic.

Titers represent the average value of obtained for two rabbit antisera raised against the D15 peptide.

Titers represent the average value obtained for two guinea pig antisera raised against the D15 peptide.

MT: not tested.

TABLE 9

RABBIT IGG ANTIBODY RESPONSE TO D15-PRP CONJUGATE

Rabbit¹#	Reactive Titer Against ²			
	PRP		rD15	
	2 doses	3 doses	2 doses	3 doses
489-1	1,600	3,200	1,600	6,400
490-1	1,600	1,600	6,400	25,600

Rabbits were immunized intramuscularly with rD15-PRP conjugates (5 to 50 μ g PRP equivalent) mixed with 3 mg ALPO, per mL, followed by two booster doses (half amount of the same immunogen) at 2 week intervals.

Reactive titres is based on PRP specific and D-15 specific ELISAs.

SUMMARY OF THE DISCLOSURE

In summary of this disclosure, the present invention provides purified and isolated nucleic acid molecules containing genes encoding the D15 outer membrane protein, the sequences of these genes and the derived amino acid sequences thereof. The invention also provides peptides corresponding to portions of the D35 outer membrane protein. In addition, the invention provides antibodies raised against D15 outer membrane protein, fragments and 10 peptides. The genes, DNA sequences, antibodies and peptides are useful for diagnosis, immunization and the generation of diagnostic and immunological reagents. Vaccines based on expressed recombinant D35, portions thereof or peptides derived from the provided sequences 15 can be prepared for prevention of H. influenzae disease. Modification are possible within the scope of the invention.

REFERENCES

O'Hagan (1992).

Ulman et al., 1993.

Berns C.A. and Thomas C.A. (1965) J. Mol. Biol. 11:476-490.

Thomas W.R. and Rossi A.A. (1986) Infect. Immun. 52:812-817.

Thomas W.R. et al. (1990) Infect. and Immun. 58:1090-1913.

Carlone G.M. et al. (1986) J. Clin. Microbiol. 24:330-331.

Smith, D.B. and Johnson K.S. (1988) Gene 67:31-40.

Harkness, R. et al. (1992) J. Bacteriol. 174:2425-2430.

Hamel et al. (1987) J. Med. Microbiol. 23:163-170.

Mills et al. (1993) Infect. Immun. 61:399-410.

Trinchieri, (1993) Immunol. Today 14:335-338.

Hope, T.P. (1986) J. Immunol. Methods 88:1-18.

Zangwill et al., 1993. MMWR 42:1-15.

Loeb et al. 1987. Infect. Immun. 55:2612-2618.

Panezutti, 1993. Infect. Immun. 61: 1867-1872.

CLAIMS

What we claim is:

- 1. A purified and isolated nucleic acid molecule, the molecule comprising at least a portion coding for a D15 outer membrane protein and having a DNA sequence selected from (a) the DNA sequence set forth in any one of Figures 1A to 1E or its complementary strand, and (b) DNA sequences which hybridize under stringent conditions to the DNA sequences defined in (a).
- 2. The molecule of claim 1 wherein said DNA sequences defined (b) have at least 90% sequence identity with the sequences defined in (a).
- 3. The molecule of claim 1 wherein said DNA sequences defined in (b) comprise the consensus sequence set forth in Figure 1F.
- 4. A recombinant plasmid adapted for transformation of a host, the recombinant plasmid comprising a plasmid vector into which has been inserted a DNA segment comprising at least an 18 bp fragment of a DNA molecule of claim 1, 2 or 3
- 5. The recombinant plasmid of claim 4 which is selected from plasmid DS-712-2-1 having ATCC accession number 75604 deposited November 4, 1993 and plasmid JB-1042-5-1 having ATCC accession number 75606 deposited November 4, 1993.
- 6. A recombinant vector adapted for transformation of a host cell, the recombinant vector comprising at least a DNA segment comprising at least an 18 bp fragment of a DNA molecule of claim 1, 2 or 3 and expression means operatively coupled to the DNA segment for expression of the gene product encoded thereby in the host cell.
- 7. The recombinant vector of claim 6 being plasmid DS-880-1-2 having ATCC accession number 75605 deposited November 4, 1993 and encoding the D15 gene product of $\underline{\text{H.}}$ influenzae SB33.

- 8. The recombinant vector of claim 6 wherein said DNA segment encodes a polypeptide of at least 6 residues.
- 9. The recombinant vector of claim 8 wherein said polypeptide is selected from those shown in Table 2.
- 10. The recombinant vector of claim 6, 7, 8 or 9 wherein said DNA segment consists of no more than the coding sequence for said D15 outer membrane protein.
- 11. The recombinant vector of claim 10, wherein the DNA segment further comprises a nucleic acid sequence encoding a leader sequence for export of said gene product from said host.
- 12. A purified and isolated protein encoded by the DNA fragment contained in the recombinant vector of claim 10 or 11.
- 13. A purified and isolated D15 outer membrane protein, or a portion thereof.
- 14. The protein of claim 13 wherein the D15 outer membrane protein is a <u>Haemophilus</u> D15 outer membrane protein.
- 15. The protein of claim 14 wherein the D15 outer membrane protein is a <u>Haemophilus influenzae</u> D15 outer membrane protein.
- 16. The protein of claim 15 wherein the <u>Haemophilus</u> influenzae is a type b <u>Haemophilus influenzae</u> strain.
- 17. The protein of claim 16 wherein the <u>Haemophilus</u> influenzae type b strain is selected from Ca, MinnA and Eagan strains.
- 18. The protein of claim 15 wherein the <u>Haemophilus</u> influenzae is a non-typeable <u>Haemophilus influenzae</u> strain.
- 19. The protein of claim 18 wherein the non-typeable Haemophilus influenzae strain is selected from PAK12085 and SB33 strains.
- 20. A synthetic peptide containing an amino acid sequence corresponding to the amino acid sequence of the protein or portion thereof claimed in any one of claims

- 11 to 19, or variant or mutant which retains immunogenicity.
- 21. The peptide of claim 20, having an amino acid sequence selected from those contained in Table 2.
- 22. An immunogenic composition, comprising a nucleic acid molecule claimed ni any one of claims 1 to 5, a protein as claimed in any one of claims 12 to 14 or a peptide as claimed in claim 20 or 21 and a physiologically-acceptable carrier therefor.
- 23. The immunogenic composition of claim 22 formulated as a vaccine for <u>in vivo</u> administration to protect against diseases caused by <u>Haemophilus</u>.
- 24. The immunogenic composition of claim 23 formulated as a microparticle preparation, capsule preparation or liposome preparation.
- 25. The immunogenic composition of claim 23 in combination with a targeting molecule for delivery to specific cells of the immune system or to mucosal surfaces.
- 26. A method for inducing protection against disease caused by <u>Haemophilus</u>, comprising administering to a subject an effective amount of the immunogenic composition of any one of claims 22 to 25 to provide protective immunity against <u>Haemophilus</u> disease.
- 27. An antiserum or an antibody specific for a protein, a peptide or an immunogenic composition of any one of claims 12 to 25.
- 28. A chimeric molecule, comprising a protein as claimed in any one of claims 12 to 19 or a peptide claimed in claims 20 to 21 linked to another polypeptide or protein or a polysaccharide.
- 29. The chimeric molecule of claim 28 wherein said another polypeptide or protein comprise a surface protein or peptide corresponding thereto from a pathogenic bacteria.

- 30. The chimeric molecule of claim 29 wherein said another polypeptide or protein comprises a P1, P2 or P6 outer membrane protein of \underline{H} . influenzae.
- 31. The chimeric molecule of claim 28 wherein said polysaccharide comprises a PRP molecule from \underline{H} . $\underline{influenzae}$.

FIG.1A.

H. influenzae b Ca strain D15 sequence

Hind III -35 GATTACGCCAAGCTTAATTTAATGATTTTACGTC <u>TATAAT</u> TTAT 10 20 30 30 60	RBS ATAGGATACAATCGATGAAAAACTTCTAATCGCAAGTTTATTATTCGGTACGAACGA	HR VAL PHE ALA ALA PRO PHE VAL ALA LYS ASP ILE ARG VAL ASP GLY VAL GIN GLY ASP L ₩ C T G T G T T T G C G C A C T T T T G T G C A A A A G A T A T T C G T G T G G T G T T C A A G G T G T T C A A G G T G T T C A A G G T G T G T G T T C A A G G T G T T C A A G G T G T T C A A G G T G T T C A A G G T G T T C A A G G T G T T C A A G G T G T T C A A G G T G T T C A A G G T G A T G G T G T T C A A G G T G A T G G T G T T C A A G G T G A T G G T G T C A A G G T G A T G G T G T C A A G A T A T T C G T G T T C A A G G T G A T G G T G T C A A G A T A T T C G T G T T C A A G G T G T C A A G A T A T T C G T G T T C A A G A T A T T C G T G T T C A A G G T G T C A A G A T A T T C G T G T T C A A G A T A T T C G T G T T C A A G A T A T T C G T G T T C A A G A T A T T C G T G T T C A A G A T A T T C G T G T T C A A G A T A T T C G T G T T C A A G A T A T T C G T G T T C A A G A T A T T C G T G T T C A A G A T A T T C G T G T T C A A G A T A T T C G T G T T C A A G A T A T T C G T G T T C A A G A T A T T C G T G T T C A A G A T A T T C G T G T T C A A G A T A T T C G T G T T C A A G A T G T T C A A G A T A T T C G T G T T C A A G A T A T T C G T G T T C A A G A T G T T C A A G A T A T T C G T G T T C A A G A T C A T C A T C A A G A T C A T C A A A A G A T C	EU GLU GLA GLA ILE ARG ALA SER LEU PRO VAL ARG ALA GLY GLA ARG VAL THR ASP ASN A TAGAACAACTTACTGTTCGTGCGGTCAGCGTGTGACTGAC	SP VAL ALA ASN ILE VAL ARG SER LEU PHE VAL SER GLY ARG PHE ASP ASP VAL LYS ALA HATGIGCTATATTCGTAAGGCCC 250 250 250 250 260 270 270 280 280 290 300	IS GIN GIJU GIX ASP VAL IEU VAL VAL VAL VAL AIA LYS SER ILE ILE SER ASP VAL L A T C A A G A A G G C G A T G T T G T T A G C G T T G T G G C T A A A T C G A T C T T C A G A T G T T A 330 330 340 350
Hind GATTACGCCAAGC 10	RBS A T A G G A T A C A A T C 70 start truncated GST/D15 —	HR VAL PHE ALA AI CTGTGTTTGCCGC	EU GLU GIN GIN II TAGAACAACAAAT	SP VAL ALA ASN II ATGTGGCTAATAT 250	IS GIN GIU GIY AS ATCAAGAAGGCGA 310

			2/82			
1 G C G 420	, L A A 480	с _С СС 540	л С 6 600	^ဗ ဗ	^ဗ ဗ ဝ	A 0
ASN A A C	VAL G T	THR LACGC 540	LEU A TTGG 600	MET ATG 66	PHE G TTCG 720	GLN I CAAA 780
ALA G C T	SER A G T	ASN A A T	LYS A A A	GIN GLU GLN AAGAACAA 650	GIN C A A	ALA 3 C A (
ASP G A T	LYS A A A I	VAL	ALA 3 C A 7	GLU 3 A A (ALA 3 C G (LYS A A A (
LEU FTA 410	ALA 3 C C 7 470	ILE A T T (530	LYS A A A (590	GLN 3 A A (GLY 3 G T G 710	ALA ; C C ? 770
ASN A A C	PHE I T T (PRO C C T A	ASP 3 A T I	LEU FTA(GLU GLY ALA GIN PHE G GAAGGTGCGCAATTCG 710	TYR ? A T G
GIN C A A	GLU 3 A A T	VAL GLU PRO ILE VAL ASN THR 1 3 T T G A A C C T A T T G T C A A T A C G C 520 530 530	ASP 3 A T (THR LEU ACATTA	PHE TTC	GLY 3 G C T
O THR GLU ALA LEU LYS GLN ASN LEU ASP ALA ASN C CACTGAAGCACTTAAACAAACTTAGATGCTAACG 390 420	LEU ILE ARG CLU LYS LEU ASN CLU PHE ALA LYS SER TTAATTCGAGAAAATTAAATGAATTGCCAAAAGT 450 460 470		LEU ILE GIN ILE ASN GLU ASP ASP LYS ALA LYS LEU A TTAATTCAAATGAAGATGATAAAGCAAAATTGG 570 580 600	ASN GLU SER VAL SER SER THR LEUGINGLU AACGAATCTGTTAGTAGCAGTACATTACAAGAA 630 640 650	LYS LEU TRP GLY ASN LYS PHE AAATTATGGGGAAATAAATTT 690 700	TYR TYR LEU ASN ASN GLY TYR ALA LYS ALA TATTATTAAATAGCTATGCCAAAGCA 750 760
LEU	LEU	ALA THR 3 C A A C A (ASN A A T	SER AGCA	ASN A A T A	ASN A T P
ALA G C A	LYS A A A	GLY ARG TYR ASN ALA THR VAL GGTCGCTATAACGCAACAGTT 510	ILE A T C /	SER AGT1	GLY ASIN GGAAA1	LEU LTAA
CLU GAA 10	GLU GAA i0	TAAC	GIN ILE CAAATC 0	VAL GTT)	TRP T G G (TYR FAT1
THR (A C T G 390	ARG G C G A G 450	TYR T A T 1 51(ILE G ATTC 570	SER VICTG	LEU T TTAT 690	TYR 1 I A T T 750
S S S	ILE A T T	ARG C G C	LEU TTA	GLU GAAA	LYS A A A	. .
ILE A T T	LEU T T A	CLY GGT	ILE A T T	ASN A A C	TRP T G G	ARG ♥ ASP CGTGA 1
SER VAL CTGTT 380	VAL G T T 440		GLU G A A 560	GLY 3 G G 620		
SERT	ASP G A T	SER AGT	ALA G C T	LYS A A G	SER T C T	SER ILE TCAATT 740
ASN	G G C	ALA G C A	ARG C G C	PHE TTC	PRO ASP SER C T G A T T C T 7 70 ST/D15	LEU GIN TGCAG
GLY G G T 370	VAL G T T 430	TYR T A T 490	ASN ASN ARG ALA GLU AATAATCGCGCTGAA 550 560	THR A C T 610	PRO C C T 670 GST/D1	LEU TTG 730
YS. ILE LYS GLY ASN SER VAL A A A T C A A G G T A A C T C T G T T 370 380	LY PHE LYS VAL GLY ASP VAL GGTTTAAAGTTGGCGATGTT 430 440	YS GLU HIS TYR ALA SER VAL A A G A G C A C T A T G C A A G T G T A 490 500	EU PRO ASN ASN ARG ALA GLU TACCAAATAATCGCGCTGAA 550 560	LA SER LEU THR PHE LYS GLY CATCATTAACTTTCAAGGG 610 620	IU LEU GLA PRO ASP SER TRP A A T T A C A A C C T G A T T C T T G G 670 end truncated GST/D15	LU LYS_ASP LEU GLN SER ILE AGAAGATTTGCAGTCAATT 730 740
ILE A T C	PHETTT	CLU GAG	PRO C C A A	SER T C A	TTA dtrun	LYS _ A A A
YS. A A A	117 G G	YS. A A	EU TAC	LA C A	LU A A en	LU A G

_			82		
ASP V GATG 840	G. G.	ARGS CGTA 960	AN ALA ILE LYS ALA LYS LEU GLY GLU ARG GLY TYR GLY S NTGCAATTAAAGCAAAACTTGGAGAACGCGGTTACGGTA 990 1020	ER ALA THR VAL ASN SER VAL PRO ASP PHE ASP ASP ALA ASN LYS THR LEU ALA ILE THR L GCGCAACGGTAAATTCAGTACCTGATTTTGATGCAAATAAACATTAGCGATAACC 1030 1040 1050 1050 1060	EU VAL VAL ASP ALA GLY ARG ARG LEU THR VAL ARG GLN LEU ARG PHE GLU GLY ASN THR V TTGTTGATGCTGGACGTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCG 1090 1100 1110 1140
ILE A T T C	GLY GGA(ARG C G C	TYR GLY FACGGT	ILE A T A 1	ASIN A T P
THRACC	ASN LEU ATCTG 890	PHE TTC(GLY	ALA 3 C G 1	GLY 3 G A P
VAL G T A 830	ASN A A T 890	THR A C T T 950	GLU ARG GLY 3 A A C G C G G T 1 1010	LEU F T A (1070	GLU 3 A A C 1130
ASN AAT	GLY GGT	ASP GATA	CLU GAA(THRACAS	ARG PHE GLU GLY GCTTTGAAGGA 1130
THR LYS VAL ACAAAAGTTA 820	ILE A T A G	EU LEU SER ALA LEU HIS LEU ASN ASP THR PHE FACTTTCAGCATTACATTTAAATGATACTTTC 930 940 950	LYS LEU GLY AAACTTGGAO	ASN LYS THR LEU AATAAACATTAG 160 1070	ARG C G C 7
LYS A A A A 820	11.E A T T A 880	LEU TTAP 940	LEU A C T T 1000	ASN A A A T 1060	ARG GIN LEU GCCAACTTC 1120
THR	ARG C G C	HIS CAT	LYS A A A	ALA G C A 1	GLN C A A
LYS	ALA GCAC	LEU	ALA G C A	ASP GAT	ARG C G C
ASP CLU LYS 3 A T G A A A A A A 810	SER TAGTO 870	A ALA A G C A T	E LYS TAAAC 990	ASP G A T 50	VAL GTT LO
T G A S		SER LTCAC 930	GLU ASN ALA ILE 3 A A A A T G C A A T T A 980	PHE 1	THR ACTG
JASN AAAT	SP LEU ACCTT	LEU A C T T	ALA I G C A	ASP : GATT	LEU 'TTA
1 <u>ге</u> G С Т А А 10	ASP TGACO	PRO LEU CTTT?	A A A 7	VAL PRO 3 T A C C T 1040	ARG ARG : G A C G T 1100
TCAG 800	TYR G T A T 860	PRO A C C T 920	GLU A G A A 980	VAL 1 G T A 1040	ARG C G A 1100
ASP VAL GIN HATGTTCAG 800	LEU GIN TYR TACAGTAT (LEU GLU	ALA ASP VAL CAGATGTAO 70	VAL ASN SER FTAAATTCAC 30	GLY GGA
THR ASI A C G G A 790	7. LEG T.T.T.	G C T	ASP A G A '	ASN AAA	ALA
S THE A A C 790	J GLY A G G T 5 850	CGAGC	T G C 970	VAL 3 G T 1 1030	VAL ASP ALA GLY STTGATGCTGGAC 1090
THR LYS	TGA	TGCCC	TATT	ALA THR CAACGC	VAL FGT
LE THR LYS THR ASP VAL GIN LEU ASN ASP GLU LYS THR LYS VAL ASN VAL THR ILE TTACTAAAACGGATGTTCAGCTAAATGATGAAAAAAAAGTTAATGTAACCATT 790 830	AL ASN GLU GLY LEU GLN TYR AS TAAATGAAGGTTTACAGTATG1 850 860	ET SER ALA GLU LEU GLU PRO LE TGTCTGCGAGCTTGAACCTTT 910 920	ER ASP ILE ALA ASP VAL GLU AS GTGATATTGCAGATGTAGAAAA 970 980	ALA	VAL rgrrg
1 E	AI T	百日	E C	展 0 0 C 0	EU TTG

AL SER ALA ASP SER THR LEU ARG GLN GLU MET ARG GLN GLN GLU GLY THR TRP TYR ASN S TTTCTGCTGATAGCACTTTACGTCAGGAAATGCGCCAACAAGAAGGAACTTGGTATAATT 1150 1150 1160 1170 1170 1180	ER GLN LEU VAL GLU LEU GLY LYS ILE ARG LEU ASP ARG THR GLY PHE PHE GLU THR VAL G CACAATTAGTTGAGTTAGGAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAACAGTCG 1210 1220 1230 1230 1260	LU ASN ARG ILE ASP PRO ILE ASN GLY SER ASN ASP GLU VAL ASP VAL VAL TYR LYS VAL L A A A A C C G A A T T G A T G G T A G T A A T G A T G A G T G G A T G T C T A T A A A G T C A A T G A T G A G T G A G T C A 1310 1320	YS GLU ARG ASN THR GLY SER ILE ASN PHE GLY ILE GLY TYR CLY THR GLU SER GLY ILE S A A G A A C G T A A C T T T G G T A T T G G T T A C G G T A C A G A G A G T G G T A T T A L 1330 1330 1340 1350 1350	ER TYR GLW ALA SER VAL LYS GLW ASP ASN PHE LEUJ GLY THR GLY ALA ALA VAL SER ILE A GTTATCAAGCAAGTGTTAAACAAGATAATTTCTTGGGAACAGGGGGGGCAGTAAGTA	LA GLY THR LXS ASN ASP TYR GLY THR SER VAL ASN LEU GLY TYR THR GLU PRO TYR PHE T CTGGTACGAAAATGATTATGGTACGAGTGTCAATTTGGGTTATACGGAGCCCTATTTA 1450 1460 1470 1480 1500	HP. LYS ASP GLY VAL SER LEU GLY GLY ASN VAL PHE PHE GLU ASN TYR ASP ASN SER LYS S CTAAAGATGGTGTAAGTCTTGGTAAATGTTTTCTTTGAAAACTACGATAACTCTAAAA 1510 1520 1560
TYR	THR V	LYS V	GLY]	SER]	TYR F ATT	SER I CTA
TRP F G G 1	GLU 3 A A A	TYR FATA	SER A G T G	VAL 3 T A A	PRO C T	ASN 1 A C T
THR A C T ' 1190	PHE T T C (1250	VAL G T A 1	GLU G A G 7 1370	ALA G C A (1430	GLU PRC GAGCC 1490	ASP G A T A 1550
GLY G G A	班 T T C	VAL G T C	THR ACA	ALA G C G	THR ACC	TYR TAC
GLU GAA	G G T	ASP G A T	aly G G T	0 0 0	TYR T A T	ASIN A A C
GLN 1 C A A 1180	THR ' A C A 1240	VAL A G T G 1300	TYR T A C 1360	THR A C A 1420	GLY G G T 1480	GLU G A A 1540
GLN C A P	ARG CGT	GLU GAA	G G T	G G A	LEU TTG	PHE TTT
ARG C G C	ASP G A T	ASP G A T	ILE A T T	LEU	ASN A A T	雅 TTC
MET A A T G	LEU TTA	ASN 1 A A T	<u>сту</u> ' с с т 50	PHE 'TTC 10	VAL GTC 70	VAL GTT 30
GLU 3 G A A 11	ARG C G C	SER AGT	PHE TTT	ASN A A T	SER A G T	ASN A A T 15
GLN C A G	ILE	GLY GGT	ASN A A C	ASP G A T	THR ACG	GLY GGA
ARG	LYS A A A	ASN A A T	ILE A T C	GIN C A A	CLY G G T	G G T
LEU TTA 1160	GLY G G A 1220	ILE A T C 1280	SER A G T 1340	LYS A A A 1400	TYR T A T 1460	LEU C T T 1520
THR	LEU	PRO C C T	CLY GGT	VAL GTT	ASP G A T	SER AGT
SER	GW GA G	ASP	TER	SER AGT	ASN A A T	VAL G T A
ASP F G A 1 1150	VAL A G T 1 1210	ARG ILE ASP PRO ILE ASN GLY CGAATTGATCCTATCAATGGTP 1270 1280	ASN F A A C 1330	ALA 1 G C A 1390	LYS 3 A A A 1450	GLY ' G G T 1510
ALA F G C	LEU ATT/	ARG	ARG	GLN C A P	THR	ASP G A T
SER	GLN A C A A	LU ASN AAAACC	GLU A G A A	TYR	LA GLY THR LYS ASN CTGGTACGAAAAATG 1450	LYS A A A
AL T	置 つ	LU A A	YS A A	E 0	LA C I	HP. C.T.

4/82

CTGATAACAAATACTACAAACTAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAG

GLN GLY

VAL.

ASP

ALA

FE

LYS

ASIN LYS TYR TYR

臤

1840

ATCACCTCTGGGTTGTATCTGCAAAGCATCTGCAGGATATGCAAATGGTTTTGGAAACA

ALA

TYR

GLY

ALA

S

ALA

ALA LYS

SEX

VAL

LEU TRP VAL

FIG.1A.(CONTINUED)

			5/82	•
雅 P T C C 1620	ж а 3.та 1680	EN G ATG 1740	3 A G 1800	/ S TT
T T T	SE IAG	AS I	ARG	រដ្ឋា ខេត
G G	ILE A T	G. G. 7	ASN A A T	PRO C C A
LEU TTA	LYS A A A	LYS A A A	LEU	ILE \ T T (
THR A C T 1610	ASN A A T 1670	PHE T T 7 1730	SER 1 G C (THR CT7
VAL 3 T T	TYR 'ATI	LYS . A A 1	ASIN ACA	VAL TTA
ASIN 1 A T (開 ここり	MET TGA	IYR A T A	ARG G A G
ER ASP THR SER SER ASN TYR LYS ARG THR THR TYR GLY SER ASN VAL THR LEU GLY PHE P GTGATACATCCTCTAACTATAAGCGTACGACTTACGGAAGTAATGTTACTTTAGGTTTCC 1570 1580 1590 1600	RO VAL ASN GLU ASN ASN SER TYR TYR VAL GLY LEU GLY HIS THR TYR ASN LYS ILE SER A CTGTAAATGAAATAACTCCTATTATGTAGGATTAGGTCATACCTATAATAAATTAGTA 1630 1640 1650 1650 1650 1660	SN PHE ALA LEU GLU TYR ASN ARG ASN LEU TYR ILE GLN SER MET LYS PHE LYS GLY ASN G ACTTTGCTCTAGAATATAACCGTAATTTATATTCAATCAA	LY ILE LYS THR ASN ASP PHE ASP PHE SER PHE GLY TRP ASN TYR ASN SER LEU ASN ARG G GCATTAAAACAAATGACTTTGATTTTCTTTTGGTTGGAACTATAACAGCCTTAATAGAG 1750 1760 1760 1770 1770 1780	LY TYR PHE PRO THR LYS GLY VAL LYS ALA SER LEU GLY GLY ARG VAL THR ILE PRO GLY S GCTATTTCCCAACTAAAGGGGTTAAAGCAAGTCTTGGTGGACGAGTTACTATTCCAGGTT 1810 1820 1830
GLY SER GAAG 1600	31.Y G T C 16	A A T	RP 1 G G A 178	EY G G T G
TYR A C G	EU (TAG	HE (LY 1 GTT	EU G
£ 5	IX 1	R 1	평 - 단 -	E C
R T GA 1590	A G (л т. АТА 710	TTT 770	A A G 830
THI YAC	VAL	150 17 1	SB T C I	ALA G C J
ARG C G I	TYR FAT	ASN 1 A T	PHE TTT	LYS A A
LYS A A G	TYR 'AT'	ARG G T 2	ASP A T T	/AL TTA
TYR ' A T 1	SER C C T 1640	ASN A C C 1700	PHE T G T 760	E G G G 1820
ASN AC1	ASN ACT	FYR A T A	SP A C T	XS (A A A G
ER ASP THR SER SER ASN TYR GTGATACATCCTCTAACTA1 1570 1580	RO VAL ASN GLU ASN ASN SER CTGTAAATGAAAATAACTCO 1630 1640	SN PHE ALA LEU GLU TYR ASN ACTTTGCTCTAGAATATAAC 1690 1700	LY ILE LYS THR ASN ASP PHE GCATTAAAACAAATGACTTT 1750 1760	HR I CTA
CI	, D, 1	A G	R A A A	A A
A T (гс <i>1</i>	LE T C T 1690	TH A A C 1750	PR C C 1810
THA C	A A A	ALA G C	LYS A A A	PHE T T (
ASP 3 A T	VAL 3 T A	PHE	ILE TT	TYR A T
G T (RO C T C	SN A C 1	LY G C A	G C T
				-

FIG. 1A. (CONTINUED)

		6/8	32		
YS ARG LEU PRO PHE TYR GIN THR TYR THR ALA GLY GLY ILE GLY SER LEU ARG GLY PHE A A G C G T T T A C C G T T C T A C T T A T A C A G C G G G G G G G G G T C G G T T A C G G G T T T T G G T T T T G T T T G C T G G T T T A C B C G G G G G G G G G G G G G G G G G	LA TYR GLY SER ILE GLY PRO ASN ALA ILE TYR ALA GLU TYR GLY ASN GLY SER GLY THR G CTTATGGTAGTATTGGACCTAACGCAATTTATGCCGAATATGGTAGTGGTAGTGGTACTG 2050 2050 2060 2070 2070	ASP VAL ILE CLY CLY ASN ALA ILE ALA THR ALA SER A GATGTGATTGGTGGTAATGCAATCGCTACAGCG 2130 2140 2150 2160	LA GLU LEU ILE VAL PRO THR PRO PHE VAL SER ASP LYS SER GIN ASN THR VAL ARG THR S CAGAGTTAATTGTGCCAACTCCATTTGTGAGCGATAAGAGCCAAAATACGGTCCGAACCT 2170 2180 2190 2220	SER VAL TRP AGN THR LYS TRP LYS SER ASP LYS AGN GLY L GTGTTTGGAATACTAAATGGAAATCAGATAAAATGGAT 2250 2280	EU GLU SER ASP VAL LEU LYS ARG LEU PRO ASP TYR GLY LYS SER SER ARG ILE ARG ALA S TAGAGAGCGATGTATTAAAAGATTGCCTGATTATGGCAAATCAAGCCGTATTCGCGCCT 2390 2330 2340
GLY 3 G T	GLY GGT?	ALA 3 C T 7	ARG THR GAACC 222	ASN AT(ARG
ARG GT(SER G T G	THR A GA G	ANL TCC	XS A A A	ILE .
LEU F T A C 2030	GLY 3 G T A 2090	ALA 3 3 C T A 2150	THR VAL ACGGTCC 2210	ASP LYS SATAAAA 2270	ARG 1 2 G T A 2330
SER I	SN SA T G	LE A	ATA 2	A B A C A	SER AI
GLY S	GLY ASN GTAAT	ALA ILE SCAATO	GIN ASN	S S	A A G
සි විවිධ	មិ ១ ។	AL. I G C	GLA C C A	LYS 3 A A	SB A T C
ILE A T 2020	TYR 1 T A 2080	ASIN 7 A A 2 2140	SER 3 A G C C 2200	LYS TRP LYS SER AAATGGAAATCA(2260	LYS 2 A A A 2320
GLY CLY ILE GTGGCATCC	ALA GLU TYR 3 C C G A A T A T C 2080	G G T	LYS A A G	LYS A A A	C G C
CLY GGT	ALA G C C	GLY GGT	ASP GATA	ASN THR AATACTA)	TYR F A T
ALA G C G C	ILE TYR .T.T.A.T.C 2070	ILE A T T	SER AGCG	ASN A A T	ASP TYR GLY LYS SER SATTATGGCAAATCAA 2320
THR AACAG	ILE T ATTT 2070	VAL ILE CLY CLY ASN GTGATTGGTGGTAATC 2130	VAL 3 3 T G A 2190	TRP 1	PRO 7 C T G 2310
TYR T A T A	ALA G C A I	ASP GAT(PHE TTTG	VAL 3 T T 1	LEU FTGC
THR ACTI	ASN A A C (SER TCT	PRO C C A 1	SER AGTO	ARG A G A 1
TYR GLN ATCAA 2000	PRO C C T 2060	SER A G T 2120	THR A C T 2180	ER LEU PHE VAL ASP ALA ALA CCTTATTGTTGATGCGGCAA 2230 2240	LYS A A A 2 2300
TYR T A T	ILE CLY LTTGGA	ILE A T A	PRO THR CCAAC1 2180	ALA ALA 3 C G G C P 2240	LEU
班 TTC1	ILE A T T	LY THR PHE LYS LYS ILE SER GTACTTTTAAGAAGATAAGT 2110 2120	GLU LEU ILE VAL AGTTAATTGTGC 2170	ASP 3 A T G	EU GLU SER ASP VAL LEU TAGAGAGCGATGTATTA? 2290
LEU PRO TACC1 1990	SER 3 A G T A 2050	PHE LYS	11.E A T T (170	LEU PHE VAL TATTTGTTG 2230	ASP C G A T C 2290
LEU TTA 1	GLY 3 G T	PHE L T T Z	LEU TA	PHE 7. T. T. C. 22	SER GCC
YS ARG AGCGT1	LA TYR GLY CTTATGGTA 20	THR	GLU A G 1	LEU TAT	ELU A G A
YS A G (C T T	G T A	LA C A G	ER J	EU (TAG

	_	_	
	Ć	_)
Į	L	L	J
	_	_)
•	2	2	•
i	_	_	•
4	Z	2	•
()
(_)
`		7	
۹ •	\ -	_	
•	•	;	
<u>ر</u>	_	_	1
L	1	-	

PHE GLN TRP GLN SER PRO ILE GLY PRO LEU VAL PHE SER TYR ALA L	TYR GLU ASN ASP ASP VAL GLU GIN PHE GLN PHE SER ILE GLY GLY S
TTCCAATGGCAATCTCCTATTGGGCCATTGGTATTCTCTTATGCCA	TATGAAAATGATGATGTCGAACAGTTCCAATTTAGTATTGGAGGTT
2360 2370 2380 2390 2400	2420 2430 2460
TRP GIN SE	ASN ASP ASI
ATGGCAATC	AAATGATGA
ER THR GLY VAL GLY PHE GLN T	YS PRO ILE LYS LYS TYR GLU A
CTACAGGTGTCGGATTCCAAT	AACCAATTAAAAAATATGAAA
2350 2360	2410 2420

	. G С С Т А А 2520
	A A C G T T C T C T 2510
	C T C A A A A A C A A C G T T C 2500 2510
	ттсатса вааст 2490
	СТТТТТТ 2480
***	AATTGAA
* *	C T A A T A 2470
田田田田田田田田田田田田田田田田田田田田田田田田田田田田田田田田田田田田田田	СТТТ (
12	5

7/82
A A T G A A 2580
АТАТТТАТСА 2570
T A A T T A A G G 2560
A C C C A T C A T T 2550
A A A A T A T T A A A 2540
T T G G G C A G A G
TTTAA

TGCTTC
СТТСА G G С T A 2630
TTGCACTTG 2620
G C T T T A G G T A 2610
A C C G C A C T T 2600
С G C A A A A G T A 2590
AAACAT

AGATCGC
A T C A C C C 2690
TATTTTCAAC 2680
A A T G C A G G T A T 2670
т G С т т т с а т т а 2660
A A G A A A A A A T 2650
СССТС

	4	Y T T A G C A G C A	0366
	AAACCTGTACTACTACAAA	TARREDIO	2750
	AACCTGTAG		7/40
	CTGAATTT	0220	00.13
	AAACTTGATG	2720	
	_	2/10	
ζ	ראאטרטפ		

8/82

77
Z
=
Z
\asymp
$\mathbf{O}_{\mathbf{i}}$
<u> </u>
\triangleleft
_
, ;
\odot
正

G T T 2820	C A G 2880
AAAA	CCAA
. G A A G C 2810	. A A A C G 2870
4 G T A G	r c a a a
AAAA	ATATI
C G T A 2	G C T G 1 2860
SATAAAATTGCTGCTCGTAAAAAAGTAGAAGCAAAAGTT 180 2790 2800 2810 2820	3 C A C C T C G C T C A A G C T G A T A T T C A A A A C G C C A A C A G A C A C A C
C T G C 2790	TACG 2850
AATTG	r c g c r
80 A T A A A	CACC1
G A T C	G A T G 284
AGTT	AAAA
: A A A A A A G A A C 2770	TAGA 2830
CAAAA	GCTT
₽ G	ອ ວ ອ

GATAAAAAA

.. 0 74/12041

FIG.1B.

DS-712-2-1 DVA, Eagan D15 sequence IS THE SEQUENCE BEING TRANSLATED

T T G 60
GCGAT
C C T G (
ACTT
AATT
A G G G A 40
TATI
A A A A 30
ССТТБ
TTAAC 20
CCTT
TTTC
10 A G C
C A G G A
AC

G C A A A T 120
T G G C C C A T C A G
CAAAAGGTGC 100
A T T T C T A T T G 90
TAAGTGGGCCA. 80
АТТАААТААТТ 70
T C J

ATTATG 6 180 88	
AATTTAGGG1 170	
АТТАСТСТА 160	
A T G G C A C T G 150	
ттаастттт 140	
G T G T A T T T T 30	
ATTGGATTG 13	

G A A G C T 240
тттаасаат g с
тсаттта g ттт 220
A G A T G G C G G 210
тасса <u>стат</u> 200
тттатттссат 190
A A

		10/8			
THR A C 420	LEU TT 480	ASP G A 540	HIS C A 600	LYS A A 660	GLY G G 720
THR ACG/	ASP G A C	ASN A A T	ALA G C G	VAL LY! GTTAA 660	ASN A A C
THR THR	GLY 3 G T	ASP 3 A C	LYS A A A (SER ASP VAL CAGATGT1	ALA CT2
THR CG1	GIN : A A G	THR ACTG	AL TGA	ER CAG	SP ATG
LEU PHE GLY THR THR THR TH TTATTCGGTACGACAACGAC 420	ARG VAL ASP GLY VAL GLN GLY ASP LES CGTGTGGATGGTGTTCAAGGTGACTT 460 470 480	VAL ARG ALA GLY GIN ARG VAL THR ASP GTTCGTGCGGTCAGCGTGTGACTGAC 510 530	VAL ALA ASN ILE VAL ARG SER LEU PHE VAL SER GLY ARG PHE ASP ASP VAL LYS TGTGGCTAATATTGTCCGCTCTTTATTCGTAAGTGGTCGATTCGATGTAAA 550 580 590	GIN GIJ GLY ASP VAL LÆJ VAL VAL SER VAL VAL ALA LYS SER ILE ILE SER ASP TCAAGAÄGGCGATGCTTGTTGTTAGCGTTGTGGCTAAATCGATCATTTCAGAT 610 620 630	ILE LYS GLY ASN SER VAL ILE PROTHR GLU ALA LEU LYS GLN ASN LEU ASP ALA ASN GL) A A T C A A A G G T A A C T C T G T T T C C C A C T G A A G C A C T T A A A A C T T A G A T G C T A A C G G 670 680 690 720
PHE TC	GLY	ARG	ASP 3 A T G	ILE T C A	ASN A A C T
LEU PTAT	ASP GLY	GIN A G C	PHE .	SER ILE ILE CGATCATTI	JIN 1
LEU TTAT	VAL 3 T G (щу 3 G T С 520	ARG C G A T 580	LYS A A A T 640	LYS GIN AAACAA 700
MET LYS LEU LEU ILE ALA SER LEU TAGGATACAATGAAAAACTTCTAATCGCAAGTTTA 370 380 390 400	ARG C G T (ALA G C C G	GLY 3 G T (ALA 3 C T 2	LEU CTTA
ALA G C A A	LYS ASP ILE AAAGATATT 450	ARG CGTG	SER AGT(VAL GTG(ALA 3 C A (
ILE A T C 390	ASP G A T 450	VAL G T T (510	VAL G T A 570	VAL G T T 6	GLU GAAA 690
LEU CTÀ	LYS A A A	PRO C C T (PEE TTC	SER AGC	THR ACTO
LEU CTT	ALA ; C A	LEU TTA	ARG SER LEU CGCTCTTTAT 560	VAL GTT?	PRO C
5 LYS A.A.A.A.C 380	VAL G T G G 10	SER LEU AGTTT?	SER T C T	VAL G T T (ILE ATT(
LYS A A A 38	PHE TTTT 440	ALA S G C A A 500	ARG S CGCT 560	ASP VAL LEU VAL VAL; ATGTGCTTGTTGTT 620	VAL ILE 3 T T A T I 680
MET A T G	PRO C C T	ARG C G A C	ILE VAL	VAL G T G	SER T C T (
T C G	ALA G C A (ILE A T C (ILE A T T	ASP 3 A T	ASN A A C T
C A A 370	ALA G C C (430	GLN C A A 2	ASN A A T 7 550	алу 3 G С (610	GLY 3 G T 1 670
ATA	PHE TTT(GLU GIN GIN ILE ARG ALA SER LEU A G A A C A A A T C C G A G C A A G T T T A 490 500	ALA G C T	GIN GLU AAGAA(ILE LYS GLY TCAAAGGT1
A G G	VAL GTG1	GLU GAA	VAL 3 T G (GIN	ILE \ T C !
H	E	A (T.	T (A A

ASN VAL THR ILE ASP

GLU LYS THR LYS VAL

GLN LEU ASN ASP

ASP

THR LYS

ς ₀	5		1/82		
LYS A A A 780		ALA G C 900		GLU C G A 1020	ILE A T 1080
VAL G T A	TR ACG	LEU	MET ATG	班 I.T.C.(1(GIN C A A A
SER A G T G	ASN A A T	ALA LYS	GIN	GLN A A	ALA C A (
LYS A A	PL TC1	M.A C A P	ilu A A C	LIA C G C	YS A A G
ALA LYS CCAAA 770	LE V TTG 830	LYS A	LEU GIN GIU GIN TACAAGAACAA 950	PHE GLU GLY ALA GLN 'TTGAAGGTGCGCAAT 1010	ALA LYS CCAAAC 1070
PHE A	O I	ASP L	J G A C 1	J G AGG	T G C
I PH A T T	PR A C C	AS) I' G A	LEI ATT	G. A. S. S. A.	TYR TATC
GLU GAA1	GLU GAA	ASP GAT	THR ACA1	PHE TTT	ASN GLY .ATGGC1 1060
ASN A A T G 760	VAL G T T 820	GLU G A A 880	SER A G T 940	LYS A A A T 1000	ASN 1 A T 1060
LEU FTAA	THR ACA(ASN A A T (SER G C J	ASN 1 A T 1	ASN A A T P
LYS LEU ASN GLU PHE ALA LYS SER VAL LYS A A A T T A A A T G A A T T T G C C A A A G T G T A A A 760 770	ALA G C A A	ILE A T C A	SER GTA	GLY 3 G A A	LEU TAA
GLU G A A A 750	GLU HIS TYR ALA SER VAL GLY ARG TYR ASN ALA THR VAL GLU PRO ILE VAL ASN THR A G A G C A C T A T G G G C G C T A T A A C G C A A C A G T T G A A C C T A T G T A T A C G 790 830	GIN ILE ASN GLU ASP ASP LYS ALA LYS CAAATCAATGAAGATGATAAAGCAAAA 870 880 890	VAL SER SER THR LEU GLN GLU GLN MET GTTAGTAGCAGTACATTACAAGAACAAATG 930 950	ე ე	_ I
ARG GAG	TYR ATA		SER V	LEU TRU PTATG 990	TYR TYF [ATTA 1050
ILE P	D I	Ri A	4	LYS LA	TT
J II AAT	ARG TCGC	F-4	<i>r</i> >	LYS 3 A A	ASP
. LEU TTTA 740	, GLY A G G 800	J ILE AAT 860		2 TRP G T G (980	ARG C G T
ASP VAL FATGTT1	SER VAL GLY GTGTAGG 800	ALA GLU ILE CTGAAAT 860	G G G	TRP T G G 98	ILE AATTC 1040
	SER AGT		SER LEU THR PHE LYS CLY ASN ATCATTAACTTTCAAGGGAAC 910 920	LEU GLA PRO ASP SER TRP TRP ATTACAACCTGATTCTTGGTGG 970 980	LYS ASP LEU GIN SER ILE ARGGAAAGATTGCAGTCAATTCGT 1030 1040
LYS VAL CLY AAGTTGGCO 730	TYR ALA A T G C A 1 790	ARG	PHE	ASP	ASP LEU GLN ATTTGCAGT 1030
VAL 3 T T (730	TYR A T (ASN A T (850	IR C T 1 910	PRO C T G 970	LEU (T. G. C. 1030
DYS A A G	GLU HIS	ASN A A T A	SER LEU THR CATTAACTI	GIN E	SP I
PHE I TTA	J D I	PRO A	A T	LEU GI	A G A
다 G T 1	GI A G A	PR A C C	SE A T C	LE A T T	LYS S A A A (
			·	. 7	O

12/82						
MET L A T 1200	SER P A G 1260	SER 7 A G 1320	LEU 1.EU 1.380	VAL C G T 1440	SER F T C 1500	
G G 7	ARG C G 1	G G T	THR ACC	THRACC	ASN AAT	
GLY GGA	ARG C G C	TYR T A C	ILE A T A	ASN A A T	TYR 'A T	
LEU CTG	PHE T C	GLY	ALA C G	GLY GA 2	RP G G 1	
ALA ARG ILE ILE GLY ASN LEU GLY GLY CACGCATTATAGGTAATCTGGGAGGTA 1180 1190	ASN ASP THR FABTECTT 1250	ARG G C G C G (LYS THR LEU A AAACATTAG 1370	ARG FHE CLU CLY ASN THR VAL. GCTTTGAAGGAAATACCGT 1430	GIN GLU GLY THR TRP TYR AAGAAGGAACTTGGTAT1 1480	
G G T	ASP GAT	GLY GLU	THR ACA	PHE I T T	GLY 3 G A	
ILE A	A A T	GLY G G A	LYS A A A	ARG CGC:	GLU 3 A A C	
ILE A T T 1180	LEU T T A A 1240	LEU C T T (1300	ASN A A T 1 1360	LEU CTT 1420	GLN C A A (1480	
ARG C G C	ALA LEU HIS SCATTACATI	ILE LYS ALA LYS ATTAAAGCAAAAC 1290	ASP ALA	GLN C A A	GIN C A A	
ALA GCA	LEU TTA	ALA G C A	ASP G A T	THR VAL ARG GIN ICTGTTCGCCAA 1410	ARG C G C	
SER F A G T G 1170	ALA A G C A 1230	LYS L A A A 1290	PHE ASP TTGATG 1350	VAL G T T 410	MET A A T G 1470	
ARG CGT 1	SER TCAC 12	ILE	PHE TTT	THR ACT 1	GLU 3 A A	
LEUCTT	LEU CTT	ALA G C A	PRO ASP CTGATT	LEU	GIN C A G (
ASP GACC	LEU TTAC 0	ASN A A T G O	PRO C C T	ARG C G T	ARG	
TYR A TATG 1160	PRO C C T 1 122(ASP VAL GLU ATGTAGAA 1280	SER VAL ICAGTAC 1340	ARG ARG LEU ' CGACGTTTAA 1400	ASP SER THR LEU ARG GLN GLU MET SATAGCACTTTACGTCAGGAAATG 1450 1470	
GLY LEU GLN 3 G T T T A C A G T 1150	GLU GAA	VAL 3 T A	SER CCA	ALA CLY SCTGGA(班 C T J	
LEU I'T A	LEU	ASP 3 A T	ASN ATT	ALA CTC	SER GCA	
GLY G G T 1150	ALA GLU LEU GLU SCGAGCTTGAAC 1210	ALA 3 C A C 1270	VAL 3 T A A 1330	ASP 3 A T G 1390	ASP ; A T A 1450	
GAA G	ALA G C C	ASP ILE ATATTO	ALA THR VAL	VAL FTT	ALA 3 C T G	
ASN GLU GLY LEU GLN TYR ASP LEU ARG SER ALA ARG ILE ILE GLY ASN LEU GLY GLY MEI AAATGAAGGTTTACAGTATCACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTAT 1150 1160 1160 1160 1200	SER ALA GLU LEU GLU PRO LEU LEU SER ALA LEU HIS LEU ASN ASP THR PHE ARG ARG SER GTCTGCCGAGCTTGAACCTTTACTTTCAGCATTACATTA	ASP ILE ALA ASP VAL GLU ASN ALA ILE LYS ALA LYS LEU GLY GLU ARG GLY TYR GLY SER TGATATTGCAGATGTAGAAAATGCAATTAAAGCAAACTTGGAGAACGCGGTTACGGTAG 1270 1280 1290 1380	ALA THR VAL ASN SER VAL PRO ASP PHE ASP ASP ALA ASN LYS THR LEU ALA ILE THR LEG CGCAACGGTAAATTCAGTACCTGATTTTGATGATGCAAATAAACATTAGCGATAACCT 1330 1340 1350 1380	VAL VAL ASP ALA CLY ARG ARG LEU THR VAL ARG CLN LEU ARG PHE CLU CLY ASN THR VAL TGTTGTTGATGCTGGACGTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCGT 1390 1400 1410 1410 1410	SER ALA ASP SER THR LEU ARG GLN GLU MET ARG GLN GLN GLU GLY THR TRP TYR ASN SET TTCTGCTGATAGCACTTTACGTCAGGAAATGCGCCAACAAGAAGGAACTTGGTATAATTC 1450 1450 1460 1470 1500	

_		13/8			
GLU C G A 1560	LYS : A A 1620	SER F A G 1680	ALA A G C 1740	THR 1 A C 1800	SER A A G 1860
VAL G T C (VAL G T C	ILE A T T	ILE ATAC 17	РИЕ ТТТА 18	LYS A A A A 18
THR ACAC	LYS VAL AAAGTCA	SER GLY	SER AGT	TYR l A T	SER
GLU 3 A A	TYR LAT	SER G T	VAL	PRO C C 7	ASN AAC1
ARG LEU ASP ARG THR GLY PHE PHE GLU THR VAL GLA CGCTTAGATCGTACAGGTTTCTTCGAAACAGTCGA 1530 1540 1560	GLY SER ASN ASP GLU VAL ASP VAL VAL TYR LYS VAL LYS GGTAGTAATGAGTGGATGTCGTATATAAGTCAA 1590 1620	ASN PHE GLY ILE GLY TYR GLY THR GLU SER GLY ILE SER A A C T T I G G T A C G G G I A C A G A G A G G G T A T A G 1650 1680	ASP ASN PHE LEU GLY THR GLY ALA ALA VAL SER ILE ALA GATAATTTCTTGGGAACAGGGGGGGCGCAGTAAGTATAGC 1710 1720 1730	THR SER VAL AGN LEU GLY TYR THR GLU PRO TYR PHE THI A C G A G T C A A T T T G G G T T A T A C C G A G C C C T A T T T A C L G A G C C C T A T T T A C L B 1790 1770 1800	GLY ASN VAL PHE FHE GLU ASN TYR ASP ASN SER LYS SEA GGAAATGTTTTTTGAAAACTACGATAACTCTAAAAG 1830 1840 1860
PHE TTC1	VAL GTC	THR ACA	ALA 3 C G	THR A C C	TYR FAC
THR GLY 1 C A G G T 1 1540	ASP GATG	TYR GLY 1 A C G G T 7	aly 3 G G (TYR F A T	ASN A A C 7
THR A C A 1540	VAL G T G (1600	TYR T A C 1660	LEU GLY THR GLY TGGGAACAGGG 1720	G.У G.G.Т. 1780	GLU 3 A A A 1840
ASP ARG SATCGTA	GA A	GLY GGT	G G A	LEU I I G (FHE LTTC
ASP G A T	ASP GLU GATGA <i>P</i>	ILE A T T	LEU TTG	ASN A A T	ASN VAL PHE AATGTTTCT 1830
LEU CTTAG 1530	ASN FAATG 1590	GLX F G G T 1650	PHE L T T C T 1710	VAL : G T C 1770	VAL 1 G T T 1830
ARG CGCT	SER AGTA 15	HE TTT	ASN A A T	SER AGTG 17	ASN AAT
ILE ATT (ASP G A T	THR A C G	G G A
LYS AAA 20	ASN A A T	ILE ATC 10	GIN C A A 10		
GLN LEU VAL GLU LEU GLY LYS A C A A T T A G T T G A G G A A A A 1510 1520	ASN ARG ILE ASP PRO ILE ASN A A A C C G A A T T G A T C C T A T C A A T 1570 1580	GIJU ARG ASN THR GLY SER ILE AGAACGTAACACGGGTAGTATC 1630 1640	TYR GIN ALA SER VAL LYS GIN TTATCAAGCAAGTGTTAAACAA 1690 1700	GLY THR LYS ASN ASP TYR GLY TGGTACGAAAATGATTATGGT 1750 1760	LYS ASP GLY VAL SER LEU GLY TAAAGATGGTGTAAGTCTTGGT 1810 1820
LEU	PRO CCTA	ASN THR GLY AACACGGGT1	VAL G T T	ASP G A T	SER A G T
VAL GLU ; T T G A G T 1510	ASP G A T	THR A C G	SER AGT	ASN A A T	VAL G T A i
VAL G T T 1510	ILE A T T G 1570	ASN A A C 1630	ALA G C A 1690	LYS A A A A 1750	GLY 3 G T C 1810
LEU TTA(ARG C G A	ARG	GIN C A A (THR ACGI	ASP
GLN	ASN A A C (GLU 3 A A C	TYR ' A T (GLY THR LYS ASN G T A C G A A A A A T C 1750	LYS A A G
A (A A	A G	T.	T G	TA

		14/				
PRO C C C 1920	ASN r A A 1980	GLY r G G 2040	GLY G G 2100	SER 7 T C 2160	ASP G A 2220	U LYS C A A 2280
路 T T C (15	SER A G T	ASN AAT(ARG A G A	G G T	ARG A G A	ASN A A C
GLY 3 G T	ILE A T T	GLY 3 G T	ASN A A T	PRO C C A	ASP 3 A C	GLY 3 G A
LEU GLY FTAGGT1)	LYS A A A	LYS	LEU	ILE \ T T (LEU	PHE TT(
ASN VAL THR LE NATGTTACTTT 1910	TYR TYR VAL GLY LEU GLY HIS THR TYR ASN LYS ILE SER TATTATGTAGGATTAGGTCATACCTATAATAAATTAGT1 0 1970 1950 1960	RG ASN LEU TYR ILE GLN SER MET LYS PHE LYS GLY ASN GL) GTAATTTATATATTCAATGAAATTTAAAGGTAATGG 2010 2010 2020	ASN TYR ASN SER LEU ASN ARG GLY AACTATAACAGCCTTAATAGAGG 2080 2090 2100	PHE PRO THR LYS GLY VAL LYS ALA SER LEU GLY CLY ARG VAL THR ILE PRO GLY FTCCCAACTAAAGGGTTAAAGCAAGTCTTGGTGGACGAGTTACTATTCCAGGT 2110 2120 2130 2130	ASP VAL GLN GLY PHE TYR PRO LEU ASP ARG 3 A T G T A C A G G G T T T C T A C C C A T T A G A C A G G G T T C T A C C C A T T A G A C A G A C	LEU TRP VAL VAL SER ALA LYS ALA SER ALA GLY TYR ALA ASN GLY PHE GLY ASN LY. TTTGGGTTGTATCTGCAAAGCATCTGCAGGATATGCAAATGGTTTTGGAAACAA 2230 2230 2240 2250
WAL	TYR I A T	LYS A A A	ASN A A C	VAL G T T	TYR TAC	ASN A A T
ASN A A T (III A C C T	MET A T G	TYR	ARG C G A (雅 T T C '	ALA 3 C A
TYR GLY SER PACGGAAGTA	HIS CAT 1	SER T C A 2	ASN A A C ' 2080	G G A (2140	GLY G G T ' 2200	TYR T A T (2260
GLY G G A	G G T	GLN CAA	PHE GLY TRP FTTGGTTGG1	GLY GGT	GLN C A G	GLY G G A
TYR TAC	LEU TTA	ILE A T T	GLY GGT	LEUCTT	VAL G T A	ALA G C A
THR 3	GLY A G G A 1950	TYR A T A T 2010	ME エエエ 2070	SER A A G T 2130	ASP G A T 2190	SER A T C T 2250
ARG THR	VAL GTA	LEU TTA	SER TCTT 20	ALA G C A	ALA GCAC 21	ALA G C A
ARG C G T	TYR T A T	ASN A A T	ASP PHE ASP PHE 3 A C T T T G A T T T T 2060	LYS A A A	SER AGTG	LYS A A A
LYS A A G C	TYR TAT	ARG C G T 10	ASP G A T	VAL G T T	LEU CTA	ALA G C A 10
ASN TYR I 1 A C T A T A 1880	ASN SER 1 NACTCCT 1940	ASN AACC AACC 2000	PHE T T T 206	стх v сссс 2120.	LYS A A A 218	SER A TCTG 2240
ASN A A C	A A C	TYR T A T	ASP G A C	LYS A A A	TYR T A C	VAL G T A
SER TCTA	ASN AATA	ALA LEU GLU TYR ASN ARG ASN LEU TYR ILE GLN CTCTAGAATATAACCGTAATTTATATATTCAA 1990 2010	ASN A A T (THR A C T	ASN LYS TYR TYR LYS LEU A A C A A A T A C T A C A A C T A A 2170 2180	VAL G T T
SER T C C 1 1870	GLU G A A 3 1930	LEU CTA 1990	LYS THR AAACAI 2050	PRO C C A 2110	LYS A A A 2170	TRP T G G 2230
THRACAT	ASN A A T C	ALA G C T	LYS A A A	雅 T T C	ASN A A C	LEU CTC
ASP THR SER SER ASN TYR LYS ARG THR THR TYR GLY SER ASN VAL THR LEU GLY PHE PR TGATACATCCTCTAACTATAAGCGTACGACTTACGGAAGTAATGTTACTTTAGGTTTCCC 1870 1870 1880 1890 1920	VAL ASN GLU ASN ASN SER TYR TYR VAL GLY LEU GLY HIS THR TYR ASN LYS ILE SER AST TGTAAATGAAATAACTCCTATTATGTAGGATTAGGTCATACCTATAATAAATTAGTAA 1980 - 1930 - 1940 - 1950 - 1960 - 1960 - 1970 - 1980	PHE ALA LEU GLU TYR ASN AL CTTTGCTCTAGAATATAACCO 1990 2000	ILE LYS THR ASN ASP PHE ASP PHE SER PHE GLY TRP ASN TYR ASN SER LEU ASN ARG GLA CATTAAAACAAATGACTTTGATTTTTTTGGTTGGAACTATAACAGCCTTAATAGAGG 2050 2060 2060 2070	TYR PHE PRO THR LYS GLY VAL LYS ALA SER LEU GLY GLY ARG VAL THR ILE PRO GLY SET CTATTTCCCAACTAAAGGGGTTAAAGCAAGTCTTGGTGGACGAGTTACTATTCCAGGTTC 2110 2120 2130 2130	ASP ASN LYS TYR TYR LYS LEU SER ALA ASP VAL GLN GLY PHE TYR PRO LEU ASP ARG ASI TGATAACAAATACTACAAACTAAGTGCAGATGTACAGGGTTTCTACCCCATTAGACAGAGA 2170 2180 2190	HIS LEU TRP VAL VAL SER A TCACCTCTGGGTTGTATCTG 2240

_		15/			
ALA P G C 2340	GLY 1 G G 2400	ALA 3 G C 2460	SER 7 T C 2520	LEU A T T 2580	SER T C 2640
PE TT1	THR ACT	SER AGCO	ARG THR	ASN GLY AATGGAT 25	ALA G C C
ARG GLY Gregri	SER CLY GTGGT	ALA G C T	ARG C G A	ASN A A T	ARG
ARG C G T 0	SER A G T	THR ACAG	VAL 3 T C	LYS	ILE . T T (
SER LEU P CATTAC 2330	TYR GLY ASN GLY S ATGGTAATGGTA 2380 2390	GLY GLY ASN ALA ILE ALA THR ALA SER ALA GGTGGTAATGCAATCGCTACAGCTAGCGC 2450 2440 2460	VAL SER ASPLYS SER GIN ASN THR VAL ARG THR SEG GTGAGCGATAAGAGCCAAAATACGGTCCGAACCTC 2490 2520	R VAL TRP ASN THR LYS TRP LYS SER ASP LYS ASN GLY LES TGTTTGGAATACTAAATGGAAATCAGATAAAATGGATT 2550 2550 2560	ARG ILE CGTATTC 2630
SER T C A	ASN A A T	ILE A T C	ASN AATA	SER TCAG	SER AGC(
GLY GGTT	GLY GGT	ASN ALA ILE NATGCAATCO 2440	SER GIN IGCCAAI 2500	TRP LYS CGGAAAT 2560	LYS SER AAATCAA 2620
ILE A T C 2320	TYR T A T 2380	ASN A A T 2440	SER A G C 2500	TRP r g g z 2560	LYS A A A C 2620
G G C 1	GLU GAA	G G T	LYS A A G A	LYS A A A T	GLY 3 G C 1
CLY GGT(TYR ALA GLU FTATGCCGAA 2370	VAL ILE CLY FTGATTGGTC 2430	SER ASP AGCGATA 190	THR ACT/	ASP TYR GLY SATTATGGC?
ALA A G C G G 2310	TYR F T A T 2370	ILE 3 A T T 2430	SER A G C 6	ASN 3 A A T 7 2550	ASP P G A T 7 2610
THR ACAC 23	ILE ATT 2	WAL GTG 2	VAL GTGA 24	TRP FGG	PRO C C T (
TYR GIN THR TYR 'ATCAAACTTATA 2300	ALA G C A	ASP G A T	PRO PHE CCATITO)	ALA SER VAL ; C A A G T G T T ' 2540	LEU FTG(
THR ACT 00	ASN A A C	SER TCT	PRO C C A O	SER AGT(ARG A G A S
GLN 1 C A A A 2 2300	PRO ASN CCTAA(2360	SER S AGTT 2420	THR FACT C 2480	ALA : 3 C A A 2540	LYS 7 A A A A A 2
TYR T A T	ILE GLY	LYS LYS ILE A G A A G A T A A	ILE VAL PRO TTGTGCCA 2470	VAL ASP ALA FTGATGCGC 2530	LEU PTA
PHE TTC1	ILE A T T	LYS A A G	VAL 3 T G	ASP 3 A T (VAL
PRO C C G 7 2290	SER A G T A 2350	LYS A A G 1 2410	ILE A T T (2470	VAL 3 T T G 2530	ASP VAL LEU SATGTATTAA 2590
LEU TTA(GLY 3 G T	PHE LTT	LEU LTA1	PHE TTC	SER 1 G C G
ARG LEU PRO PHE TYR GLN THR TYR THR ALA GLY GLY ILE GLY SER LEU ARG GLY PHE ALA GCGTTTACCGTTCTATCAAACTTATACAGCGGGTGGCATCGGTTCATTACGTGGTTTTGC 2290 2330 2340	TYR GLY SER ILE GLY PRO ASN ALA ILE TYR ALA GLU TYR GLY ASN GLY SER GLY THR GL) TTATGGTAGTATTGGACCTAACGCAATTTATGCCGAATATGGTAGTGGTAGTGGTACTGG 2350 2350 2360 2370 2370	THR PHE LYS LYS ILE SER SER ASP VAL ILE TACTTTTAGGAGATAAGTTCTGATGTGATT 2410 2430	GLU LEU ILE VAL PRO THR PRO AGAGTTAATTGTGCCAACTCC 2470 2480	LEU PHE VAL ASP ALA ALA SER CTTATTGTTGATGCGGCAAG 2530 2540	GIU SER ASP VAL LEU LYS ARG LEU PRO ASP TYR GLY LYS SER SER ARG ILE ARG ALA SER A G A G A G C G A T G T A A A A A A A G A T T G C C T G A T T A T G G C A A A T C A A G C C G T A T T C G C G C C T C 2590 2630 2640

_
Ш
\supseteq
\underline{Z}
F
Z
0
S
<u>·</u>
\mathbf{m}
\Box
正

LYS	SER
A A	'T C
2700	2760
ALA G C C	G G T
IRP GIN SER PRO ILE GLY PRO LEU VAL PHE SER TYR ALA LYS	ASN ASP ASP VAL GIJJ GIN PHE GIN PHE SER ILE GLY GLY SEA
GGCAATCTCCTATTGGCCATTGGTATTCTCTTATGCCAA	A T G A T G A T G T T C C A A T T T A G T A T T G G A G G T T C
2670 2680 2690 2700	2730 2740 2750
SER	ILE
TCT	ATT
30	50
PHE	SER
TTCT	A G T A
2690	2750
VAL	PHE
G T A	T T T
LEU	GIN
TTG	C A A
PRO	PHE
C C A	T T C
2680	2740
G G G	GIN CAG
ILE	GU
A T T	GAA
PRO	VAL
C C T	G T C
2670	2730
SER	ASP
TCT	GAT
GLN	ASP
C A A	GAT
F 0	7 K 0
GIN C A A T 2660	G A A A A 2720
PHE	TYR
TTC	T A T
GLY	LYS
G G A	A A A
VAL	LYS
G T C	A A A
2650	2710
G G T	ILE A T T
THR GLY VAL GLY PHE GIN	PRO ILE LYS LYS TYR GLU
TACAGGTGTCGGATTCCAA	ACCAATTAAAAAATATGAA
2650 266	2710 272

16/82	
ст G С С Т А А Т	C A A A T G A A A
2820	2880
АСААСGТТСТ	G G A T A T T T A T
2810	2870
TTCTTCAGAACTCAAAAACACGTTCTCTGCCTAA	TAAACCCATCATTAATTAAGGATATTATCAAATGAA
2790 2800 2810 282	2850 2850 2850
тттсттсатся	ттааасссат 2850
ТСТААТАААТТ G ААСТТТ 2780	A A T T G G G.C A G A G A A A A T A 9
2770	2830 2840
тстаатаа; 2770	A A T T G G G C 7

AACCGCACIIIAGGIAIIGCACIIGCIIIGCIICAGGCTATGC	2930		
I C C A C I I G	2920		
TRIBORIT TOO.	2910		
	2900		
	2890		

TTCC 2940

1 C1/CA33/003

300

290

GTTATCGAATTGGCGCAGCACTGTTATTAAGCTTAACGGTGTTTGCATTATTAATGATT

270

260

250

	1	7	/	8	2	
,						(

DS-691-1-5 DWA, Minn A D15 sequence IS THE SEQUENCE BEING TRANSLATED

FIG.1C.

ATTGGATTGGTATTTTTAAGTTTTATGGCACTGATTA 120 90 120	AATTTATTTCCATTACCAGTATTAGATGGCGGTCATTTAG 150 160 170 180	G A G C G G G T G C A A A G C A T C T 230
TAAGTTTTATGGC 00 110	A T T A G A T G G C G G 170	G A G C G G G T G C A 230
TAAGT 00	A T T A	G A G
T T T T T 1	C C A G T 7	G T T T C T 220
G A T T G G T G T A T 90	ТАТТГССАТТА 150	GTTAAAGGAAACCTGTTTCTGAGCG 210
C A A A T 80	TATG 140	A A G C T 200
TGCTGG	тстааатттассс 130	ттттттаасаат 6 б
	G G C A C A T C A G C A A A T 70 80	TGCTGGCACATCAGCAAAT 70 80 TGTAAATTTAGGGATTATG 140

			18/82			
LEU TTA	360	VAL G T G 420	GLY G G T 480	GLY ARG GGTCGA 540	ALA LYS G C T A A A 600	LYS A A A 660
SER A G T		ARG C G T	ALA G C C		ALA G C T	LEU
ALA G C A		PHE VAL ALA LYS ASP ILE TTTGTGGCAAAAGATATT 400	ARG CGTC	LEU PHE VAL SER TTATTCGTAAGT 530	VAL GTGO	ALA G C A
ILE A T C	350	ASP G A T 410	VAL G T T 470	VAL G T A 530	VAL G T T (590	LYS GLY ASN SER VAL ILE PRO THR GLU A A A G G T A A C T C T G T T A T T C C C A C T G A A 630 640 650
LEU C T A		LYS A A A	PRO C C T	PHE TTC	SER A G C	THRACT
LEUCTT		ALA G C A	LEU TTA	LEU TTA	VAL 3 T T	PRO C C C
LYS A A A	0	VAL GTG	ALA SER GCAAGT 460	ARG SER CGCTCT 520	VAL G T T (ILE ATT 10
LYS LYS A A A A A A	340	PHE (TTTG 400	ALA S G C A A 460	ARG 8 CGCT 520	LEU VAL CTTGTT 580	VAL GTTA 640
MET A T G		PRO C C T	ARG C G A	VAL GTC	VAL G T G	SER T C T
T C G		ALA G C A	ILE A T C	ILE A T T	ASP GAT	ASN A A C
CAA	330	ALA ALA G C C G C A 390	GLN C A A 450	ASN ILE AATATT 510	ал ССС 570	GLX G G T 630
TAĞGATACAATCG		PHETT	GIN C A A	ALA G C T	GLU GAA	LYS A A A
A G G		VAL G T G	I GLU AGAA	VAL TGTG	GLN C A A	S ILE AATC
	320	~ E	LEU T T A 440		,, E+	
T T A		THR ACG	ASP GACT	ASN AATG	ALA G C G	VAL GTT?
AAT		THR THR ICGACAA	GLN GLY CAAGGT(ASP GACA	LYS A A A	ASP GAT(
TAT	310	THR TACG 370	VAL GIN GLY ASP LEU GTTCAAGGTGACTT 430 440	, THR G A C T G 490	ASP ASP VAL LYS ALA HIS GATGATGTGAAAGCGCA 550 560	: SER TTCA(
G T C	33	GLY GGTA	VAL G T T C 430	VAL GTGA	ASP GAT 59	ILE ATT 61
ТТТТАССТСТАТААТТАТА		LEU PHE GLY THR THR THR THR TTATTCGGTACGACAACGAC 370	aly G G T	GIN ARG VAL THR ASP ASN ASI CAGCGTGTGACTGACAATGA 490 500	ASP GAT(SER ILE ILE SER ASP VAL LYS TCGATCATTTCAGATGTTAA 610 620
TTT		LEU TTA	ASP G A T	GIN C A G (PHE T T C	SER T C G

		19/82	?		
ASN	VAL	GLU	SER	LYS	ASN
A A T	G T T	G A A	A G T	A A A	A A T
720	780	840	900	960	1020
LEU T T A	THR ACA	ASN A A T	SER SER AGCAGT 900	ASN A A T	ASN A A T
LYS A A A	ALA G C A	ILE A T C	SER A G T	G G A	LEU
ARG GLU LYS LEU ASN CGAGAAAAATTAAAT 710	TYR ALA SER VAL GLY ARG TYR ASN ALA THR VAL TATGCAAGTGTCGCTATAACGCAACAGTT 750 770 760 780	ASN ARG ALA GLU ILE LEU ILE GLN ILE A A T C G C T G A A A T T T A A T T C A A A T C 810 830	VAL G T T 890	LEU TRP GLY ASN LYS TTATGGGGAAATAAA 950	SYS ASP LEU GLA SER ILE ARG ASP TYR TYR LEU ASN ASN A A G A T T G C A G T C C A A T T C G T G A T T A T T T A A A T A A T 990 1020
ARG	TYR	ILE	SER	LEU	TYR
C G A (T A T	A T T	T C T	TTA	TAT
	ARG	LEU	GLU	LYS	ASP
	C G C	TTA	GAA	A A A	G A T
LEU TTA 10	G G T	ILE ATT 0	ASN A A C	TRP TRP GGTGG 940	ARG C G T
VAL LEU ILE	VAL (GU]	GLY 1	SER TRP TRP LYS	11.E
GTTTTAATT	G T A G	GAAA	GGGA	TCTTGGTGGAAA	
700	760	820	880	940	
ASP	SER	ALA	LYS	SER	SER
G A T	A G T	G C T	A A G	T C T	T C A
C C C	ALA G C A	ARG C G C	LEU THR PHE LYS GLY ASN TTAACTTTCAAGGGAAC 870 880	ASP GAT	GLN CAG
VAL	TYR	ASN	THR	PRO	LEU
G T T	T A T	A A T	A C T	C C T	T T G
690	750	810	870	930	990
LYS	HIS	ASN	LEU	GIN	ASP
A A A	CAC'	AAT	TTA	C A A	G A T
PHETTT	GU GAG	PRO C C A	SER TCA	LEU TTA	
GLY	LYS	LEU	ALA	GLU	GU
GGGG	A A A	C T A	G C A	GAA	GAG
680	740	800	860	920	980
ASN	VAL	ASN THR	LEU	MET	PHE
A A C	G T A I	AATACG		A T G	TTC
GIN ASN LEU ASP ALA ASN GLY PHE	GLU PHE ALA LYS SER VAL LYS GLU	GLU PRO ILE VAL ASN THR LEU PRO ASN	ASP ASP LYS ALA LYS LEU ALA SER	THR LEU GLN GLU GLN MET GLU LEU	PHE GLU GLY ALA GLN PHE GLU
CAAAACTTAGATGCTAACGGGTTT	GAATTTGCCAAAAGTGTAAAAGAG	GAACCTATTGTCAATACGCTACCAAAT	GATGATAAGCAAAATTGGCATCA	ACATTACAAGAACAAATGGAATTA	TTTGAAGGTGCGCAATTCGAGA
670 680	730	790 800	850 860	910 920	970
л ASP	LYS	ILE VAL	3 ALA	и спл	7 ALA
А G A T G	CAAAA	\TTGTC?	AGCA	А G A A	TGCG
670	730	790	850	910	970
LEU TTA (ALA G C C A 730	ILE A T T	LYS A A A C 850	GIN C A A 9	G G T
ASN	PHE	PRO	ASP	LEU	GLU
A A C	TTT(C C T A	GATI	TTA(GAA
GIN	GLU	GLU	ASP	THR	PHE
C A A	GAA	GAA(G A T	ACA	TTT(

1350

1340

1330

FIG.1C.(CONTINUED)

		20/82			
LYS A A A 1080	ILE A T T 1140	LEU T T A 1200	LEU CTT 1260	ASN A A T 1320	LEU CTT
THR A C A	ARG ILE CGCATT 1140	HIS C A T	LYS A A A	ALA G C A	GIN C A A
LYS THR ASP VAL GIN LEU ASN ASP GLU LYS THR LYS A A A A C G G A T G T T C A G C T A A A T G A T G A A A A A A C A A A A A A A A A A A A	ALA G C A	LEU P.T.A	ALA 3 C A	GLY GLU ARG GLY TYR GLY SER ALA THR VAL ASN SER VAL PRO ASP PHE ASP ALA ASN GGAGAACGCGGTACGGTAGCGCAACGGTAATTCAGTACCTGATTTTGATGCAAAT 1270 1280 1290 1390	ASP ALA GLY ARG ARG LEU THR VAL ARG GLN LEU GATGCTGGACGTTTAACTGTTCGCCAACTT
GLU F G A A 7	SER ' A G T (ALA 1 G C A 1 1190	LYS 7 A A A (ASP 1 G A T 1310	LEU THR VAL PTAACTGTT
ASP GAT(ARG CGTA	SER TCAG	ILE ATT? 12	PHE TTT 1	THRACT
ASN A A T (LEU CTT	LEU	ALA G C A	ASP GAT1	LEU TTA
LEU CTA1	ASP GAC(LEU TTA(ASN A A T (PRO CCT(ARG CGTT
VAL GIN I STTCAGC 1060	TYR 7 TATG 1120	FRO 1 CCTT 1180	GLU 1 GAAA 1240	VAL G T A C 1300	ARG C G A C
VAL GTT	GIN CAG1	GU GAA	VAL G T A (SER T C A	GLY G G A C
ASP GATG	LEU TTAC	LEU CTT	ASP G A T	ASN AAT7	ALA G C T G
THR A C G (1050	GLY G G T 1110	GLU G A G 1170	ALA G C A 1230	VAL G T A 1290	ASP GAT(
LYS A A A A	GU GY GAAGGT 1110	ALA G C C	ILE A T T	THR ACG	VAL GTT
THR ACTA	AAAT	SER GTCT	ASP GAT	ALA G C A I	VAL G T T (
11.E \ A T T 1040		MET A T G 1160	SER ' A G T 1220	SER ' A G C 1280	
GLN C A A A 10	ASP G A T	G G T A	ARG CGTA	GLY GGT?	THR ACC
ALA G C A C	ILE A T T (GLY G G A (PHE ARG	TYR TACG	ALA ILE THR LEU SCGATAACCCT
LYS AAAC 30	祖 A C C <i>i</i> 90	LEU CTG(PHE TTC LO	ARG GLY : G C G G T 1 1270	ALA G C G
ALA 1 G C C A 1030	VAL 1 G T A A 1090	ASN I AATC 1150	ASP THR 13.1.0 12.10	ARG (C G C G 1270	LEU TTA(
GLY TYR ALA LYS ALA GLN ILE GGCTATGCCAAAAT 1030 1040	VAL ASN VAL THR ILE ASP VAI GTTAATGTAACCATTGATGT 1090	ILE GLY ASN LEU GLY GLY MET ATAGGTAATCTGGGAGGTAT 1150 1160		GIU GAAC	LYS THR LEU ALA ILE THR LEE AAAACATTAGCGATAACCCT
C C C	VAL G T T	ILE A T A (ASN AAT C	GLY GGAO	LYS A A A

SER VAL ASN LEU

GLY

ASP

ASN

LYS

景

GEV

ALA

ILE

贸

1680

1620

1610

GLY

E

ASN PHE

VAL LYS GLN ASP

段

ALA

S. S.

TYR

SER

E

EZ G

1650

FIG.1C.(CONTINUED)

		21/82	
GLN C A A 1440	THR A C A 1500	GLY PHE PHE GLU THR VAL GLU ASN ARG ILE ASP PRO ILE ASN GLY SER ASN ASP GLU VAL NO GENTATION OF THE GENT OF THE GE	TYR T A C
GIN	ARG	GLU	GLY
C A A	C G T	GAA	GGT
ARG	ASP	ASP	ILE
C G C	GAT	G A T	A T T
MET	LEU	ASN	GLY
A A T G	CTTA	1 A A T	G G T
1430	1490	1550	610
GLU GAA 1	ARG CGC 1	SER AGT	PHE GU. TTTGG
GIN	ILE	G G T	ASN
C A G	A T T		A A C
ARG	LYS	ASN	ILE
C G T	A A A	A A T	A T C
O	0	O	0
LEU AR	GLY I	ILE. 1	SER]
TTACG	GGAA	A T C A	A G T A
1420	1480	1540	1600
ARG PHE GLU GLY ASN THR VAL SER ALA ASP SER THR LEU ARG GIN GLU MET ARG GIN GLU	GLU GLY THR TRP TYR ASN SER GLN LEU VAL GLU LEU GLY LYS ILE ARG LEU ASP ARG THR	ASN ARG ILE ASP PRO ILE ASN GLY SER ASN ASP GLU VAL	ASP VAL VAL TYR LYS VAL LYS GLU ARG ASN THR GLY SER ILE ASN PHE GLY ILE GLY TYR
CGCTTTGAAGGAAATACCGTTTCTGCTGATAGCACTTTACGTCAGGAAATGCGCCAACAA	GAAGGAACTTGGTATAATTCACAATTAGTTGAGTTAGGAAAAATTCGCTTAGATCGTACA	A A C C G A A T T C A A T G G T A G T A A T G A T G A G T G	GATGTCGTATATAAGTCAAAGAACGTAACAGGGTAGTATCAACTTTGGTATTGGTTAC
1390 1430 1410 1410	1450 1460 1470 1500	1530 1540 1550	1570 1580 1590 1600
SER	GU	ASP	THR
AGC	GAG	G A T	ACG
ASP	VAL	ILE	ASN
G A T	G T T	A T T	A A C
1410	1470	1530	1590
ALA	LEU	ARG	ARG
G C T	TTA	C G A	C G T
SER	GIN	ASN	GU
TCT	C A A	AAC	GAA
VAL	SER	GLU	LYS
GTT	7 T C A	: G A A	: A A A
1400	1460	1520	1580
THR	ASN	VAL	WAL
ACC	A A T	G T C	G T C
ASN	TYR	THR	LYS
A A T	T A T	A C A	A A A
ARG PHE CLU CLY ASN THR VAL	GLU GLY THR TRP TYR ASN SER	PHE GLU THR VAL GLU	ASP VAL VAL TYR LYS VAL ATGTCGTATATAAAGTCA 1570
GCTTTGAAGGAAATACCGT	BAAGTTGGTATAATTC1	TCGAAACAGTCGA <i>1</i>	
1390 1400	1450 1460	1510 1520	
GLU GAA	THR 7 ACTT 1450	PHE (エエCG 1510	VAL G T A T 1570
PHETT	GLY	PHE	WAL
	G G A	TTC1	G T C
ARG	GLU	GLY	ASP
C G C	GAA	GGT1	G A T

GGTACAGAGAGTGTTAGTTATCAAGCAAGTGTTAACAAGATAATTTCTTGGGAACA GGGGGCGCAGTATAGCTGGTACGAAAATGATTATGGTACGAGTGTCAATTGG 段 ALA VAL 1570 THR GLU ALA SUBSTITUTE SHEET

		22/82			
GLU G A A 1800	SER A G T 1860	HIS C A T C 1920	SER T C A 1980	ASN A A C 2040	GLY G G A
PHE TTT	GLY G G A	GLY GGT	GIN C A A	TRP T G G	G G T
PHE TTC	TYR GLY SER TACGGAAGT 1860	LEU TTA	ILE A T T	C C T	LEU
VAL : G T T 1790	ТНК 3 А С Т '	GLY A G G A 1910	TYR 1 T A T 1970	PHE TTT 2030	SER A A G T (2090
ASN AAT	THR ACGA	VAL GTA	LEU T.T.?	SER T C T 1	ALA G C A A
GLY G G A ,	ARG C G T	TYR T A T	ASN A A T	PHE TTT	LYS A A A
GLY G G T 30	LYS A A G (TYR T A T 00	ARG C G T	ASP GAT	VAL G T T
LEU (CTTG	ASN TYR LYS ARG THR THR A A C T A T A A G C G T A C G A C T 1840 1850	SER 1 T C C T 1900	ASN 7 A A C C 1960	PHE 7 TTTG 2020	GLY 1 GGGG 2080
LYS ASP GLY VAL SER LEU GLY GLY ASN VAL A A A G A T G T C T A G T C T T G G T G G A A A T G T T 1770 . 1780 . 1790	ASN A A C	GLU ASN ASN SER TYR TYR VAL GLY LEU GLY GAAAATAACTCCTATTATGTAGGATTAGGT 1890 1900 1910	TYR T A T	LYS THR ASN ASP PHE ASP PHE SER PHE CLY TRP ASN A A A A C A A A T G A C T T T T G G T T G G A A C 2010 2020 2030	TYR PHE PRO THR LYS GLY VAL LYS ALA SER CTATTTCCCAACTAAAGGGGTTAAAGGT 2070 2080 2090
VAL G T A	SER SER T C C T C T 1830	ASN A A T	GLU GAA	ASN A A T (THRACT
GLY ' G G T 1770	SER T C C 1830	G A A .	LEU C T A 1950	THR A C A 2010	PRO C C A 2070
ASP 1 G A T	THR 'ACA	VAL ASN GTAAAT	ALA GCT		PHE TTC
LYS	ASP G A T	VAL GTA	PHETT	(ILE CATT	TYR T A T
THR LACT 1760	SER A A G T 1820	PRO C C C T 1880	ASN 7 A A C 1940	ASN GLY ATGCC 2000	ARG GLY 1 G A G G C 2060
PHE L T T T A 17	LYS LAAAA 18	PHE 18	SER PAGTA	ASN A T	ARG A G A
TYR	SER	CLY AGGT7	ASN LYS ILE AATAAATT 1930	LYS GLY AAAGGT!	ASN AATA
JU PRO A G C C C 1750	3P ASN 1 T A A C' 1810	THR LEU 1CTTTA(1870	N LYS	E LYS TAAA 1990	SER LEU GCCTT/ 2050
THR GLU 1750	ASP CGAT 181	THR IAC	ASN FAA7	PHE A T T T A 1990	SER
	ASN TYR ASP ASN SER LYS SEA A A C T A C G A T A A C T C T A A A A G 1810 1820	ASN VAL THR LEU GLY PHE PRO A A T G T T A C T T T A G G T T T C C C 1870 1880	THR TYR ASN LYS ILE SER ASN ACCTATAATAAATTAGTAA 1930 1940	MET LYS PHE LYS GLY ASN GLY A T G A A T T T A A A G G T A A T G G 1990 2000	TYR ASN SER LEU ASN ARG GLY TATAACAGCCTTAATAGAGG 2050
TYR TAT	ASN A A C 1	ASN A A T	THR A C (MET A T G	TYR T A T A

GLY

GLY

田田

M.

ASP

SER

段

HE

LYS

盟

GLY

胃

GLY

SER

GLY

ASN

2310

2300

2290

2330

GGTAATGGTAGTGGTACTTTTTAAGAAGATAAGTTCTGATGTGTTGTGTGGTGGTA

GCAATCGCTACAGCTAGCGCAGAGTTAATTGTGCCAACTCCATTTGTGAGCGATAAGAGC

2430

SE

LYS

ASP

SE

W

出

PRO

景

PR0

VAL

ILE

E

GE

ALA

SER

ALA

ALA THR

ILE

ALA

23/82

FIG.1C.(CONTINUED)

		23102	
GLY G G T 2160	TYR T A T 2220	ILE A T C C 2280 C	TYR T A T
GLN C A G	GLY GGA	G G C	GLU GAA
VAL G T A	ALA G C A	G G T	ALA G C C
ASP A G A T 2150	SER A T C T 2 2210	A G C G 2270	TYR T A T
ALA G C A	ALA G C A 2	THR ACA 2	ILE A T T
SER A G T	LYS A A A A	TYR	ALA G C A
LEU CTA.	ALA G C A	THR ACTO	ASN A A C (
PRO GLY SER ASP ASN LYS TYR LYS LEU SER ALA ASP VAL GLN GLY CAGGTTCTGATAACAAATACTACAAACTAAGTGCAGATGTACAGGG 2120 2130 2130 2140	PHIS LEU TRP VAL VAL SER ALA LYS ALA SER ALA GLY TYR TCACCTCTGGGTTGTATCTGCAAAAGCATCTGCAGGATAT 2190 2220	ARG LEU PRO PHE TYR GIN THR TYR THR ALA GLY GLY ILE 3 G T T T A C C G G T G G C A T C C T A T A C A G C G G G G G C A T C A A C T T A T A C A G C G G G G C A T C A A C T T A T A C A G C G G G G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A G C G G T G C A T C A A C T T A T A C A C T T A T A	TYR GLY SER ILE GLY PRO ASN ALA ILE TYR ALA GLU TYR FATGGTAGTATTGGACCTAACGCAATTTATGCCGAATA
TYR I A C	VAL 3 T A	TYR F A T (GLY 3 G A (
TYR TAC	VAL G T T	RE TTC'	ILE A T T (
LYS A A A 2 2130	TRP T G G 2190	PRO C C G 2 2250	SER A G T
ASN A A C	LEU	LEU TTA	GLY G G T
ASP G A T	HIS C A C	ARG C G T	TYR
SER 'TCT 2120	ASP 1 G A T 2180	LYS : A A G 2240	ALA G C T
GLY GGT	TYR PRO LEU ASP ARG ASP FACCCATTAGACAGAGA1 2170 2180	ASN AAC 2	PHE TTT
PRO C C A	ASP G A C	GLY GGA	GLY GGT
ILE ATT 10	LEU TTA 70	PHE T T T 10	ARG C G T
THR 1 ACTA 2110	PRO I C C A T 2170	GLY FG G T T 2230	T T A C
ARG VAL THR ILE PRO GLY SER ASP ASN LYS TYR TYR LYS LEU SER ALA ASP VAL GLN GLY CGAGTTACTATTCCAGGTTCTGATAACAAATACTACAAACTAAGTGCAGATGTACAGGGT 2110 2120 2130 2130 2140	PHE TYR PRO LEU ASP ARG ASE TTCTACCCATTAGACAGAGA 2170 2180	ALA ASN GLY PHE GLY ASN LYS ARG LEU PRO PHE TYR GLN THR TYR THR ALA GLY GLY ILE GCAAATGGTTTTGGAAACAAGCGTTTACCGTTCTATCAAACTTATACAGCGGGTGGCATC 2230 2240 2250 2260	GLY SER LEU ARG GLY PHE ALA TYR GLY SER ILE GLY PRO ASN ALA ILE TYR ALA GLU TYR GGTTCATTACGTGGTTTTGCTTATGGTATTGGACCTAACGCAATTTATGCCGAATAT
ARG C G A	PHE TTCT	ALA G C A	GLY GGT

FIG. 1C. (CONTINUED)

•		24/82		
TRP T G G 2520	LYS A A A 2580	PRO C C A 2640	PHE T T C 2700	A A A 2760
LYS A A A	GLY LYS GGCAAA 2580	GLY PRO GGGCCA 2640	GIN C A G	C II C
PHE VAL ASP ALA ALA SER VAL TRP ASN THR LYS TTTGTTGATGCGCAAGTGTTTGGAATACTAAA 2490 2500 2510	PRO ASP TYR C C T G A T T A T 2570	ILE A T T	GLU ASN ASP ASP VAL GLU GIN PHE GAAAATGATGATGTCGAACAGTTC 2680 2680 2700	PHE *** *** TTCTAATAAATTGAACTTTTTTTCATCAGAACTC 2730 2740 2750
ASN 5 A A T 7 2510	ASP ' G A T 2570	РКО 'ССТ 1 2630	VAL 'GTC(1 T C A 2750
TRP TGG		SER TCT(ASP GAT(T C A
VAL GTT1	LEU TTG	GIN C A A	ASP GAT	T C T
SER AGT(LEU LYS ARG TTAAAAAGA 2560	TRP T G G (ASN AAT(T T T 10
AIA : G C A A 2500	LYS	GIN C A A T 2620	GLU 7 GAAA 2680	C T T T 2740
ALA G C G G	LEUTTA	PHE TTC	TYR T A T	GAA
ASP GAT(VAL G T A	GLY GGA	LYS LYS A A A A A A A '	ATT
VAL G T T 2490	ASP G A T 2550	VAL G T C 2610		*** T A A 2730
PHETTT	SER A G C	G G T	ILE A T T	*** T A A
LEU	I GLU SER AGAGAG	THR ACA	PRO C C A	PHE TTC
SER T C C 2480		SER 7 T C T 2600	ALA LYS. CCAAA 2660	SER TCT 2720
THR AACC 2	GLY F G G A T 25	ALA GCCT 26	ALA GCC	GLY GLY SER 3GAGGTTC' 2720
ARG	ASN A A A	ARG FCGCG	TYR	GLY G G A
THR VAL CGGTCO 2470	ASP LYS ASN HATAAAATG 2530	G ILE ; T A T T (2590	E SER Стст1 2650	R 11.E ; T A T T G 2710
THR FAC(ASP AGA7	ARG C G T A 2590	PHE \ T T C 1 2650	SER 1 A G T <i>P</i> 2710
GIN ASN THR VAL ARG THR SET CAAAATACGGTCCGAACCTC 2470 2480	LYS SER ASP LYS ASN GLY LE A A A T C A G A T A A A A A T G G A T T 2530 2540	SER SER ARG ILE ARG ALA SER TCAAGCCGTATTCGCGCCTC 2590 2600	LEU VAL PHE SER TYR ALA LYS. PRO ILE TTGGTATTCTCTTATGCCAAACCAATT 2650 2660	GLN PHE SER ILE GLY GLY SER CAATTTAGTATTGGAGGTTCT 2710 2720
GIN C A A A	LYS A A A 1	SER T C A	LEU TTG(GIN C A A T

2920

2910

2900

2890

AATTA	2820
ACCCATCATTA	2810
AAATATTAAA	2800
TGGGCAGAGA	2790
CTAATTTAAT	2780
AACAACGTTCTGC	2770

GCGGGTTATANTT
TAATGCGGGT
\TTGCTTTCATTAAT
GCTGAAGAAAAA
ATGCTTCC
TTGCTTCAGGCT

TNCAAGGCNAAGG 2950

FIG. 1D

SB33 D15

TRANSLATED	
BEING TRA	
SEQUENCE 1	
IS THE SE	
ĭ	(

	•	26/82		
ттастт	C T G G C	G T A A A T	T T T T T A	T A T C G A
60	120	180	240	
G A A A A T A T T A G G G A A A T T	G C A A A A G G T G	T T G A T T A G T	САТТТАGТТ	А G C А T C T G T
40 50	110	170	230	290
A A A A T A 40	CTATT	T G G C A	G C G G T	T G C A A
TAACCTT	GGGCCAATTT 100	TTAAGTTTTA 160	C A G T A T T A G A T G 210	C T G A G C G G G
TTTCCCTTT 30	атттаа G т G	T G T A T T T T T T T T T T T T T T T T	САТТАССА G 210	A A A A C C T G T T T 270
A G G A C A G C	тсаттааата	G G A T T G G	ттаттс	G A A G C T G T T A A A G G A A 1
20	80	140	200	250
G G C A T T G A A A A A A C	ACTGGCGATTTG	GCATCAGCAAATATT	TTAGGGATTATGAAT	A C A A T G G A A G C T
10	70	130	190	250

Ŕ

VAL ALA LYS

VAL

SER

M

VAL

E

VAL

ASP

GLY

GEO

GLN

LYS

W

650

TGTGAAAGCGCATCAAGAAGGCGATGTTGTTGTTAGCGTTGTGGCTAAATCGATCAT

900

FIG.1D.(CONTINUED)

		271	82	
G T 360	GLY G G 420	VAL G T 480	VAL G T 540	ASP G A
T A C	PHE TTC	GLY GGT0	ARG VAJ CGTGT 540	ASP G A T
T T T	LEU PTA'	ASP 3 A T (PHE LTC(
A T T 0	LEU FTA?	VAL 3 T G C	GLY 3 G T (ARG CGA 1
а Т G А 350	SER I AGTT 410	ARG (C T G 470	ALA (; C C G 530	GLY 3 G T C
TAI	ALA ; C A ?	ILE LTTO	ARG ALA GLY : G T G C C G T (SER 1 G T G
AT	ILE	ASP 3 A T A	VAL	VAL 3 T A A
340	LEU TAA	LYS A A A G 460	PRO VAL CTGTTC 520	PHE TCG
СТТААС G G T G T T T G C A T T T A A T G A T T T T T A C G T 330 340 350 360	MET LYS LEU LEU ILEU ILE ALA SER LEU LEU PHE TCGATGAAAAACTTCTAATCGCAAGTTTATTATTC 390 400 410	ALA LYS ASP ILE ARG VAL ASP GLY GCAAAAGATATTCGTGTGGATGGT 460 470	LEU 'TAC	ARG SER LEU PHE VAL SER GLY ARG PHE ASP CGCTCTTTATTCGTAAGTGGTCGATTCGAT
F G T J	LYS LYS 1 A A A A A C 190	VAL 3 T G G	ALA SER SCAAGTT 310	SER
C G G 7	LYS A A A 1 390	РРО РИЕ VAL ССТТТТСТG 450	ALA G C A 1 510	ARG CGC1
TAA	MET A T G	PRO C C T	ARG C G A G	VAL 3 T C C
GCT	I C G Z	ALA G C A (ILE A T C (ILE VAL ATTGTC
		_	GLN C A A I	
гатт 320	A T A C 380	PHE 7	GIN (ALA S C T A
ТСТ	A G G	VAL 3 T G T	GLU 3 A A (VAL 3 T G (
CAC	TAT	THRACTO	LEU FTA(ASP 3 A T G
C A G 310	T T A '	THR A C G A 430	ASP 3 A C 1	ASN A A T C
9 ລ ອ	AAT	THR ACAA	aly 3 G T (ASP 3 A C I
ATTGGCGCAGCACTGTTATTAA 310 320	СТАТААТТТАТАТА G G A T A C A A 370	THR THR THR THR VAL PHE ALA TACGACAGACTGTGTTGCC 430 440	GIN GLY ASP LEU GLU GIN GIN ILE ARG ALA SER LEU PRO VAL ARG ALA GLY GIN TCAAGGTGACTTAGAACAACAATCCGAGCAAGTTTACCTGTTCGTGCCGGTCAG 490 530	THR ASP ASN ASP VAL ALA ASN GACTGACAATGATGTGGCTAAT
	SH	SSTITLITE S	HEET	

_		28/8			
LEU 720	ALA G C 780	ILE A T 840	LYS A A 900	GLN CA 960	GLY 1 G G 1020
ASN A A C	PHETTT	PRO C C T	ASP LY GATAA 900	LEU GL TTACA 960	GLU GLY GAAGG
GIN CAAA	GU GAA	GU GAA	ASP 3 A T	THR A C A	PHE T T
LYS A A A (ASN A A T	VAL 3 T T	GLU 3 A A (SER	LYS
LEU LYS GIN ASN CTAAAACAAAAC 710	ARG GLU LXS LEU ASN GLU CGAGAAAATTAAATGAA 760	ARG TYR ASN ALA THR VAL CGCTATAACGCAACCGTT 820	ASN GLU ASP AATGAAGAT 890	SER SER THR AGCAGTACA 950	TRP LYS LEU TRP GLY ASN LYS PHE TGGAAATAAATT 1000 1010
ALA G C A	LYS A A A	ALA G C A	ILE A T C	SER AGT	GLY 3 G A 1
GU GAA	GLU GAA	ASN A A C	LEU HE GIN HE TTAATTCAAATC 880	VAL GTT?	TRP FGGO
PRO C C T 700	ARG C G A (TYR T A T 820	TLE A T T 880	SER T C T (LEU FTA 1 1000
PRO C C A	ILE A T T	ARG C G C	LEU TTA	GU GAA'	LYS A A A ?
STAACTCTATTATTCCACCT	LEU TTA	VAL GLY GTAGGT 810	GU ILE GAAATT 870	GLY ASN GLU GGGAACGAA 930	TRP TGG/
ILE A T T 690	1LE A T T 750	VAL G T A (810	G. A. A. 870	GLY G G G 930	TRP T G G '
SER TCT	ASP G A T	SER AGT	ALA G C T	LYS A A G	SER
ASN A A C	C C C	ALA G C A	ARG C G C	PHE	ASP G A T
3 GLY A G G T 680	LYS VAL AAGTT 740	TYR T A T	ASN A A T O	THR ACTO	0 [
LYS A A A 68	LYS AAA 74	HIS 7 CACT 800	ASN 1 A A T A 860	ALA SER LEU THR CATCATTAAC1 920	GLN]
ILE A T C A	PHE TTT	GWGW	PRO C C A A	SER T C A	LEU FTA(
LYS A A A A	ASN GLY 1 A C G G G 7 730	LYS A A A G	LEU CTAC	ALA 3 C A	GLU 3 A A T
VAL G T T A 670	ASN AAC 730	VAL G T A A 790	THR A C G (850	LEU F T G (910	MET A T G (970
ASP GAT	ALA G C T.A	SER AGTG	ASN AATA	ALA LYS LEU CCAAATTGO	GLN CAAA
SER ASP VAL LYS ILE LYS G TTCAGATGTTAAAATCAAAG 670 680	ASP ALA ASN GLY PHE LYS VAL GLY AGATGCTAACGGGTTTAAAGTTGG 730	GLN SER VAL LYS GLU HIS TYR CCAAAGTGTAAAAGAGCACTAT 790 800	VAL ASN THR LEU PRO ASN ASN TGTCAATACGCTACCAAATAAT 850	ALA LYS LEU ALA SER LEU THR PHE LYS AGCCAATCATTAACTTTCAAG 910 920	GLU GIN MET GLU LÆU GIN PRAGGAATTACAACC 970 980

		_				

LYS

ASIN

ALA

ASP

ASP

出

ASP

PRO

VAL

SES

ASIN

W

開

景

ASIN

GLY.

TYR

GLY

1330

1270

1370

TTCCGCCGTAGTGTATTGCAGATGTAGAAAATGCAATTAAAAGCAAATTGGGGAACG

1290

GEU

GLY

LYS LEU

ALA

LYS

ILE

ALA

ASI

CEU

VAL

ASP

ALA

ILE

ASP

SE

ARG

1230

1220

1210

1250

1310

1300

A G G T T A C G G T A A C A G T A A A T T C T G T A C C T G A T T T T G A C G A T G C A A A T A A A A C A T T

FIG.1D.(CONTINUED)

		29/82	
ALA	VAL	ASN	THR
r G C	! G T	A A	
1080	1140	1200	
TYR	ASN	α.Υ	ASP
TAT	AAT	G G Т	G A T
GLY	VAL	ILE	ASN
3 G C	3 T T	A T A	A A T
ASN	LYS	ILE	LEU
A A T	A A A O	A T T	I T A
ASN P A A T A 1070	THR LY. ACAAA 1130	ARG C G C J 119	HIS C A T
GIN ALA ILE ARG ASP TYR TYR LEU ASN ASN GLY TYR ALA CAGGCAATTCGTGATTATTTAAATAATGGCTATGC 1050 1060 1060 1060	ASP VAL GIN LEU ASN ASP GLU LYS THR LYS VAL ASN VAL GATGTTCAGCTAAATGATGAAAAAAAAAAGTTAATGT 1110 1120 1130	IEU GIN TYR ASP LEU ARG SER ALA ARG ILE ILE GLY ASN TA ACAGTA GEGA CGCA TA TA TA GGTA A GETA TA TA GGTA A GETA A G	LEU GLU PRO LEU LEU SER ALA LEU HIS LEU ASN ASP THR CTTGAACCTTTACTTTCAGCATTACATTAAATGATAC
TYR	GLU	SER	ALA
	GAA	A G T	G C A
TYR T A T 1060	ASP (T G A T G 1120	ARG C G T 1180	SER T C A
ASP	ASN	LEU	LEU
G A T	A A T	CTT	
ARG	LEU	ASP	LEU
C G T	CTA	G A C	T T A
ILE A A T T 1050	GLN F C A G 1110	TYR T A T 1170	PRO C C T
ALA	VAL	GIN	GLU
G C A	GTT	C A G	GAA
LEU TTG 10	ALA G C G 30	GLY G G T 50	GLU GAG
ASP 1	LYS P	GLU G	SER ALA GLU
G A T T	A A A G	GAAG	
1040	1100	1160	
LYS	THR	ASN	SER
A A A	ACT	A A T	
CLU	ILE	VAL	MET
GAG	A T C	G T A	A T G
PHE T T C 1030	GLN C A A 1090	ASP G A T 1150	GLY G G T
ALA GIN PHE GLU LYS ASP LEU CGCAATTCGAGAAAGATTT 1030 1040	LYS ALA GLN ILE THR LYS ALA A A A C C A C A A A T C A C T A A A G C (THR ILE ASP VAL ASN GLU GLY CCATTGATGTAAATGAAGG: 1150 1160	LEU GLY GLY MET TGGGAGGTATGT
ALA GLN PHE GLU LYS ASP LEU TGCGCAATTCGAGAAAGATTTG 1030 1040	LYS ALA GLA ILE THR LYS ALA CAAAGCACAAATCACTAAAGCG 1090 1100	THR ILE ASP VAL ASN GLU GLY A A C C A T T G A T G T A A A T G A A G G T 1150 1160	LEU GLY GLY MET SER ALA GLU TCTGGGAGGTATGTCTGCCGAG

		30/82			
GLU 'G A 1440	THR A C 1500	PHE 7. T. T. 1560	VAL : G T 1620	GLU . G A 1680	ALA G C 1740
PHETTT	GLY G G A	PHETTC	WAL GTC	THR ACA	ALA G C G
ARG	GLU GAA	G G T	ASP G A T	GLY GGT	G G G
	GIN GIN SACAAC 1490	THR ACA 0	VAL G T G 0	TYR TACG 0	THR ACAG
GIN LEU CAACTT 1430	GIN C CAAC 1490	ARG 1 CGTA 1550	GU V GAAG 1610	GLY 1 GGTT 1670	GLY 1 GGAA 1730
HIS CAC	ARG C G C	ASP GAT	ASP G A T	ILE A T T	LEU TTG
VAL GTTO	GIN GLU MET ARG GIN GIN GLU CAGGAAATGCGCCAACAAGAA 1480 1480 1490	LEU TTA	ASN AAT	GLY G G T	PHE TTC
THR A C T (1420	GLU GAA1 1480	ARG C G C T 1540	SER A G C 1600	PHE TTT 1660	ASN A A T '
LEU TTA	GIN C A G	ILE A T T	G G T	ASN A A C	ASP G A T
ARG C G T	ARG C G T	LYS A A A	ASN AAT	ILE A T C	GIN C A A
SP ALA GLY ARG ARG LEU THR VAL A T G C T G A C G A C G T T T A A C T G T T 1410 1420	LEU 7 T T A 1470	LEUGLY LYS ILE ARG LEU ASPARG THR GLY PHE PHE TTAGGAAAATTCGCTTAGATCGTACAGGTTTCTT 1530 1540 1560	ILE ASP PRO ILE ASN GLY SER ASN ASP GLU VAL ASP VAL VAL TTGATCCTATCAATGGTAGCAATGATGAAGTGGATGTCGT 1590 1600 1600	SER 1 A G T 1650	VAL LYS GIN ASP ASN PHE LEU GLY THR GLY ALA ALA GTCAAACAAGATAATTTCTTGGGAACAGGGGGGG 1710 1720 1730
GLY G G A	THR ACT	LEU TTA	PRO C C T	G G T	WAL GTC 1
ALA G C T	SER AGT	alu G A G	ASP G A T	THR ACG	SER A G T
	ALA ASP ; C T G A T 1460	LEU VAL PTAGTT(1520	ARG ILE : G A A T T 1580	G ASN 3 T A A C 7	N ALA N A G C A i 1700
VAL GTT 140	ALA G C T 14	LEU TTA	ARG CGA 15	ARG C G T 164	TYR GIN FATCAA(
VAL GTTG	SER TCTO	GIN CAAT	ASN AACO	GAA	TYR T A T
PHE TTT	THR VAL ACCGTT 1450	SER TCA(G A A	LYS A A A	ILE SER \TTAGT1 1690
THR A C C 1	THR A C C 1450	ASN A A T T 1510	THR VAL GLU ACAGTTGAA1 1570	LYS VAL LYS A A G T C A A A O 1630	ILE A T T 1690
ILE A T A A	ASN A A T	TYR TAT?	THR ACA	LYS A A A	GLY
ALA ILE THR PHE VAL VAL A AGCGATAACCTTTGTTGTTG 1390 1400	GLY ASN THR VAL SER ALA ASP SER THR LEU A G G A A A T A C C G T T T C T G C T G A T A G T A C T T T A 1470 1470	TRP TYR ASN SER GIN LEU VAL GIU TTGGTATAATTCACAATTAGTTGAG 1510 1520	GLU THR VAL GLU ASN ARG 1 CGAAACAGTTGAAAACCGAA 1570 1580	TYR LYS VAL LYS GLU ARG ASN THR GLY SER II.E ASN PHE GLY II.E GLY TYR A T A T A A G A A G G T A G G G T A G T A C T T T G G T A T A C A C G G G T A G T A C T T T G G T A T T G G T T A C A L 650 1630 1640 1670	SER GLY ILE SER TYR GLN ALA GAGTGGTATTAGTTATCAAGCA 1690 1700

GLU	ASP	THR 1920	ASN
: G A	G A		'A A
1800	1860		1980
THR ACC	TYR TAC	VAL GTT 1	TYR PAT
TYR	ASN	ASN	THR
FAT	A A C	A A T	ACC
GLY 3 G T O	GLU GAA O	SER AGT	HIS CAT 70
LEU G	PHE G	GLY S	GLY Н
FTGG	TTTG	GGAA	GGCC
1790	1850	1910	1970
ASN	PHE	TYR	LEU
A A T	TTC'	T A T	TTA
VAL	VAL	THR	GLY
3 T C	G T T	ACT	GGA
SER	ASN	THR	VAL
A G T (A A T	A C G	G T A
1780	1840	1900	1960
THR LYS ASN ASP TYR GLY THR SER VAL ASN LEU GLY TYR THR GLU	ASP GLY VAL SER LEU GLY GLY ASN VAL PHE PHE GLU ASN TYR ASP	ER SER ASN TYR LYS ARG THR THR TYR GLY SER ASN VAL THR	ASN GLU ASN ASN SER TYR TYR VAL GLY LEU GLY HIS THR TYR AST
ACGAAAAATGATTATGGTACGAGTGTCAATTTGGGTTATACCGA	RATGGTGTAAGTCTTGGTGGAAATGTTTTCTTTGAAAACTACGA	CCTCTAACTATAAGCGTACGACTTATGGAAGTAATGTTAC	NATGAAAATAACTCCTATTATGTAGGATTAGGCCATACCTATAA
1760 1770 1780 1780	1820 1830 1840 1850 1860	1890 1900 1900	1940 1950 1950 1960
G G T	G G T	LYS A A G	TYR T A T
TYR	LEU	TYR	SER
F T A T	PCTT	: T A T	T C C
1770	1830	1890	1950
ASP	SER	ASIN	ASN
GAT	AGT	A A C	A A C
ASN	VAL	SER SER	ASN
A A T	G T A		A A T
LYS	GLY	R SER	N GLU
A A A	G G T	A T C C	TGAA
50	20	1880	1940
THR I ACGA 1760	ASP G G A T G 1820	THR ACA 188	ASN A A T
ALA GLY	LYS	ASP	VAL
3 C T G G T 1	A A A (GAT	GTA
ALA	THR	SER	PRO
G C T		AGT(CCTC
ILE A T A C 1750	PHE T T T 1	LYS A A A A 1870	PHE TTCC 1930
SER	TYR	SER	G G T
AGTA	T A T	TCT/	
VAL SER ILE ALA CLY THR LYS ASN ASP TYR CLY THR SER VAL ASN LEU CLY TYR THR CLA A G T A A G T A T A C G A A A A A T G A T T A T G G T T G G G G	PRO TYR PHE THR LYS ASP GLY VAL SER LEU GLY GLY ASN VAL PHE PHE GLU ASN TYR ASI G C C C T A T T T T A C T A A G A T G T G T G G A C T A T G T T T T T G A A A C T A C G A A T G T T T T T G A A A A C T A C G A A A T G T T T T T T G A A A A C T A C G A A A T G T T T T T T G A A A A C T A C G A A A T G T T T T T T T G A A A A C T A C G A A A T G T T T T T T T G A A A A C T A C G A A A T G T T T T T T T G A A A A C T A C G A A A T G T T T T T T T T G A A A A C T A C G A A A T G T T T T T T T T T G A A A C T A C G A A A T G T T T T T T T G A A A A C T A C G A A A T G T T T T T T T G A A A A C T A C T A C G A A A T G T T T T T T T G A A A A C T A C G A A A T G T T T T T G A A A A C T A C	ASN SER LYS SER ASP THR S TAACTCTAAAAGTGATACAT 1870 1880	LEU GLY PHE PRO VAL ASN G TTTAGGTTTCCCTGTAAATG 1930 1940
	D. 100		- lede

31/82

出	ΤΤ	2040	SER	A G	2100
LYS	AAA		ASIN	AAC	
MET	A T G		TYR	TAT	
SER	r C A	0	ASIN	AAC	0
CIN	AA	2030	TRP	r G G	2090
ILE	ATT(GLY	GGT	
TYR	r A T i		阳	r T T	
TEN.	AGAATATAACCGTAATTTATATATTCAATCAATGAAATT	2020	SES	A A A T G A C T T T G A T T T T T T T G G T T G G A A C T A T A A C A G	2080
ASN	AATI		铝	PTI	
ARG	GT		ASP	3 A T	
ASIN	AACO	2010	出	r T T (2070
TYR	FAT!	8	ASP	GAC	2
ern ern	3 A A I		ASIN	A A T	
nen Ten	TA	0	强	Ú	6
ALA	CTC	2000	LYS	AAA	206
出	r T T (田田	ATTA	
ASN	AACT		GLY	3 G C 1	
SES	A G T A	1990	ASIN	A A T (2050
到日	ATT?		GLY	3 G T	
IXS	TAAATTAGTAACTTIGCT		FXS	TAAAGGTAATGGCATTAAAA	

		32/82			
THR	PRO	. .	LEU	LYS	GLU
C A C	: C C		A T T	7 A A	A G A
2160	2220		2340	2400	2460
VAL	TYR	ASN	SER	ASN	ALA
GTT	T A C	AAT(TCA	A A T	G C A
ARG C G A	PHE TTC	TYR ALA	G G T	ASN A A T	SER A G C
GLY G G A 0	СТ. ССТ 1 0	TYR TAT 0	ILE A T T 0	GLN CAA1	ALA G C T A 0
GLY (GGTG	VAL GIN GLY	ALA GLY T	GLY]	GIN GIY GIN ASN	THR 7
	GTACAGGGT	; C A G G A T	G G C A	CAAGGTCAAAAT	ACAG
	2210	2270	2330	2390	2450
RO THR LYS GLY VAL LYS ALA SER LEU GLY GLY ARG VAL THE CAACTAAAGGGGTTAAAGCAAGTCTTGGTGGACGAGTTAC 2130 2140 2160	VAL GTA	VAL SER ALA LYS ALA SER ALA GLY TYR ALA ASN GLY GTATCTGCAAAGCATCTGCAGGATATGCAAATGG 2250 2270 2280	RO PHE TYR GIN THR TYR THR ALA GLY GLY ILE GLY SER LES CGTTCTATCAAACTTATACAGCGGGTGGCATTGGTTCATT 2310 2320 2330	TYR GIN	ALA G C T A
SER	ASP	SER	ALA	TYR	ILEATCG
AGTO	G A T	TCTG	G C G C	T A T	
ALA	ALA	ALA	THR	ILE	ALA
G C A 1	G C A (G C A 1	A C A (A T T 7	G C A 7
2140	2200	2260	2320	2380	2440
LYS	SER	LYS	TYR	ALA	ASN
AAA(AGT	A A A	TAT	G C A	A A T
VAL GTTA	LEU SER ALA ASP CTAAGTGCAGAT 2200	ALA G C A	THR ACT	ASN AACG	GLY GGT
GLY	TYR LYS	SER	GIN	GLY PRO ASN ALA ILE TYR	GLY
1 G G G C	TACAAA(1 T C T (C A A	GGGCCTAACGCAATTTAT	1 G G T (
2130	2190	2250	2310	2370 2380	2430
LYS AAAC 21	TYR TAC		TYR TAT		ILE A T T
THR	TYR	RP VAL	PHETC	ER ILE	VAL
ACTA	T A C	3GGTT		GCATT	G T G
PRO C C A A 20	LYS AAAT 30		PRO C C G T 00	SER AGCA	ASP GATG
所E I	ASN I	LEU 1	LEU F	GLY S	SER P
T T C C	AACA	CTCT	TTAC	GGTA	TCTG
2120	2180	2240	2300	2360	2420
ASN ARG GLY TYR FHE PARA TAGAGGCTATTTCC 2110	ASP	HIS	ARG	TYR	SER
	G A T	CAC	C G T	TATG	A G T
G G C 1	SER	ASP	LYS	ALA	ILE
	TCT	GAT(A A G C	G C T 1	A T A !
ARG	GLY	ARG	ASN	PHE	LYS
A G A G	G G T	A G A (A A C P	T T T	A A G 1
2110	2170	2230	2290	2350	2410
ASN A A T	PRO C C A	ASP G A C	GGA	GGT.	ASN A A T
LEU	ILE PRO GLY SER ASP ASN LYS TYR AATTCCAGGTTCTGATAACAATAC 2170 2180	LEU ASP ARG ASP HIS LEU T ATTAGACAGAGATCACCTCT 2230 2240	PHE GLY ASN LYS ARG LEU P TTTTGGAAACAAGCGTTTAC 2290 2300	ARG GLY PHE ALA TYR GLY S ACGCGGTTTTGCTTATGGTA(2350	PHE ASN LYS ILE SER SER ASP VAL ILE GLY GLY ASN ALA ILE ALA THR ALA SER ALA GLA ATTRATA A GLA ATTRATA GCGCAGA ATTRATA A GATA CTAGA A GATA GCGA A GATA A GATA CTAGA A GATA A GATA A GATA CTAGA A GATA A GATA A GATA CTAGA A GATA A
	DIIDO	TITI ITC QUE	= CT		

GLU 3 A 580	6 4	0	r-3
GLA	THR	PRO	PHE
1 G A	'A C	1 C C	: T T
2580	2640	2700	2760
LEU	SER	LYS	SER
TTA	TCT	A A A	TCT
G G A	ALA G C C	ALA G C T	GGT 1
ASN A A T	ARG C G C	TYR T A T 0	G G G (
LYS A A A A 257	THR A C T 263	SER T C T 269	ILE G ATTG 2750
ASP	ARG	PRE	SER
G A T	C G T	TTT	AGTA
SER	SER	VAL	PHE
T C A	AGC	G T A	TTT/
LYS	SER	VAL	GIN
A A A A	T C A	G T G	C A A T
2560	2620	2680	2740
TRP	LYS	PRO	PHE
T G G	A A A	C C A	T T C
LYS A A A	C C C	GLY 3 G A	VAL GLU GLN 3 T C G A A C A G 1 2730
THR	TYR	SER	GLU
A C T	T A T	A G T	: G A A
2550	2610	2670	2730
ASN	ASP	PRO	WAL
A A T	G A T	C C T	GTC
TRP	PRO	SER	ASP
T G G	C C C	T C T	GAT(
VAL	LEU	GIN	ASP
GTT	TTA	C A A	GAT
SER	ASP	TRP	ASN ASP
AGT	G A C	T G G	A A T G A 7
254	260	266	2720
ALA	LYS	GIN	GLU
G C A	A A A	C A A	GAA
ALA G C G	LEU TTG	PHE TTC	LYS TYR GLU AAATATGAA? 2710
ASP	VAL	GLY	LYS
G A T	G T C	G G A	A A A
2530	2590	2650	2710
VAL	ASN	VAL	LYS
G T T	A A T	G T C	A A A I
PHE	SER	GLY	ILE LYS LYS TYR GLU ASN ASP ASP VAL GLU GLN PHE GLN PHE SER ILE GLY GLY SER PHE A A A A A A A A A A A G A G A C A G T T C C A A T T T A G T A T T G G G G T T C T T T A A A A A T G A A A A T G A G A T G T G
A T T T	GAGC	A G G T	
	ALA ALA SER VAL TRP ASN THR LYS TRP LYS SER ASP LYS ASN GLY LEU G C G C C A A G T G C A A T G C A A A T C A G A T A A A A T G G A T T A A A A T G G A T T A A A A	LEU TTA SER TCT	PHE VAL ASP ALA ALA ALA ALA ALA ALA ALA ALA ALA TRP ASN THP LWS TRP LWS SER ASP LWS ASP LED CA ASP TYR GLY LWS SER SER SER ARG THR ARG ARG

33/82

SUBSTITUTE SHEET

1			-
Ļ			J
L	1	Ĺ	J
_	_)
_	2		2
Ē			_
_			-
4			-
C)
Ć		۰)
•		-	-
•		٠	

		34/82
A C 2880	, C T 2940	
AAA	A C	
G A 1	J G (
A T	ນຶ່	
A A 70	A T 30	
T C A	C T A 2930	
T T	TGCTTTAGGTATTGCACTTGCTTCAGGCTATGCTGCAGCT 2910 2920 2930 2940	
A T 1	. C	
A T A	E U	
G G .	'TG(
A A 2	СТ	
T	CA	
F A A	5 E	таат G са G G Т тата 2970
T T 1	T A 1	T T
C A T 2850	G G T 2910	G G T 2970
AT	T A	C A
טט	E E	T G
AAA	5 5	AA
'ТА1 2840		1 A T T 2960
A T	CA C	T T Z
A T	_ອ	E E
AA	A C	ນ ອ
A G A	T S	E E
A G A 2830	a a g 2890	а à а 2950
ပ်	CA	AA
TTGGGCAGAAAATATTAAAACCATCATTTAATTAAGGATATTATCAAATGAAAAC 2830 2830 2840 2850 2850 2860 2870 2880	TCGCCAAAGTAACCGCACT 2890 2900	аа G аааааатт G С ттттат 2950 2960
T T	O El	A A

SUBSTITUTE SHEET

JB-1042-9-4 DNA, PAK D15 IS THE SEQUENCE BEING TRANSLATED

30 40 50 50 50 FA TITAA GIGGCCAATTIAA GIGGCCCAATTIAA GIGGCCAATTIAA GIGGCCAATTIAA GIGGCCAATTIAA GIGGCCAATTIAA GIGGCCCAATTIAA GIGCCCAATTIAA GIGCCCAATTIAA GIGGCCCAATTIAA GIGGCCCAATTIAA GIGGCCCAATTIAA GIGCCCAATTIAA GIGCCAATTIAA GIGCAATTIAA GIGCAATTIAA GIGCAATTIAA GIGCAATTIAA GIGCAATTIAA GIGCAATTIAA GIGCAATTIAA GIGCAATTIAA GIGCAATTIA	80 90 100 100 100 100 100 100 100 100 100	AAATTTAGGGATTATTTATTTCCATTACCAGTATTAGATGGCGGTCATTTAGTTT 190 220 230 240	TTTAACAATGGAAGGAAAACCTGTTTCTGAGCGGGTGCAAAGCATGTTA 250 290 300	T C G A A T T G G C G C A C T G T T A T T A A G C T T A A C G G T G T T T G C A T T A A T G A T T T T T T T T T 310 330 330
--	---	--	--	---

35/82

			/82		
LEU ILE ALA SER LEU LEU P CTAATCGCAAGTTTATTAT 410	ASP G GATG 480	GIN A CAGC 540	ARG SER LEU PHE VAL SER GLY ARG PHE A CGCTCTTTATTCGTAAGTGGTCGATTCG 580 590 600	ASP VAL LYS ALA HIS GIN GLU GLY ASP VAL LEU VAL VAL SER VAL VAL VAL SER LA LYS SER I GATGTGAAAGCGCATCAAGAAGGCGATGTGCTTGTTAGCGTTGTGGCTAAATCGA 610 620 630 660	GLU ALA LEU LYS GLN A GAAGCACTTAAACAAA 710
LEU	VAL G T G	GLY G G T	ARG C G A	LYS A A A	LYS A A A
SER AGT 1	ARG C G T (ALA GLY GCTGGT	GLY 3 G T (ALA 3 C T 7	LEU
ALA 3 C A 1 410	ILE A T T (470	ARG C G T (530	SER A G T G 590	VAL 3 T G (650	ALA 3 C A C 710
IIE A T C G	ASP 3 A T 1	VAL 3 T T (VAL 3 T A 1	VAL 3 T T (GLU GAA(
LEU CTAA	LYS A A A (PRO VAL CCTGTT	PHE LTC(SER AGCO	加るよう
LEU CTTC 400	ALA G C A 7	LEU T T A (520	LEU T T A T 580	VAL G T T 1 640	PRO THR CCCACT 700
LYS A A A	GLY THR THR THR VAL PHE ALA ALA PRO PHE VAL ALA LYS ASP ILE GGTACGACAAGGACTGTGTTTGCCGCACCTTTTGTGGCAAAGATATT 430 440 450 450	SER A G T	SER TCT1	VAL GTT(ILE A T T (
LYS A A A	PHETT	ALA SER GCAAGT	ARG C G C :	LEU	VAL GTT
MET ATACAATCGATG 390	A C C T 450	: ARG C C G A 510	ASN ILE VAL AATATTGTCO 570	. VAL TGTG 630	SER TCT
T C G 35	ALA G C A	ILE ATC 51	ILE ATT 57	ASP G A T G 630	ASN S AACT 690
CAA	ALA G C C	GIN C A A	ASN A A T	O D D	G G T
A T A	PHETT	GLN C A A	ALA G C T	GD GAAG	LYS A A A
A G G 380	VAL G T G 440	GLU G A A 500	VAL G T G 560	GLN C A A 620	ILE A T C 680
T A T	THRACTO	LEU TTA	ASP GAT(HIS CAT(LYS A A A
T T A	THR ACGA	ASP G A C 7	ASN A A T C	LYS ALA A A A G C G C	VAL GTT1
A A T 370	THR THR	GLY G G T 490	ASP G A C 1 550	LYS A A A 610	ASP G A T G 670
T A T	THR A C G	GIN C A A	THRACTO	ASP VAL	SER TCA(
ACGTCTATAATTTATATAGG 370	GLY GGTP	LY VAL CHN CHY ASP LEU CHU CHN CHN ILE ARG GTGTTCAAGGTGACTTAGAACAACAATCCGA 490 500 510	RG VAL THR ASP ASN ASP VAL ALA GTGTGACAATGATGTGGCT 550 560	ASP G A T	ILE SER ASP VAL LYS ILE LYS GLY ASN SER VAL ILE PRO THR ATTTCAGATGTTAAAATCAAAGGTAACTCTGTTATTCCCACT 670 680 690
A C	出 T C (LY G T	RG G T	SP A T (T C

		3	37/82			
J P A T 780	J P A C 840	~		G	T T 30	A A 10
GLI GA	GIU P GAAC 840	ASP A G A T G 900	THR L ACAT 960	PHE G TTTG 1020	GLY T GGCT 1080	VAL A G T T A 1140
ASN A A T	VAL G T T	GLU GAA	SER AGT	LYS A A A	ASN A A T	LYS A A A (
LYS LEU AAATTA 770	THR VAL ACCGT7	ASIV A A T	SER AGC	GLY ASN LYS GAAATAAA 1010	ASN A A T	THR A C A
LYS A A A 770	ALA ' G C A A 830	ILE A T C 7 890	SER A G T A 950	GLY G G A 1010	LEU T T A A 1070	GLU LYS THR ;AAAAACA 1130
GU GAAA	ASN AACG	GLN C A A	VAL G T T	TRP T G G	TYR T A T	GLU GAAA
ARG C G A (ARG TYR :GCTAT:	ILE A T T	SER TCT0	LYS LEU TRP AAATTATGGO	ASP TYR 3 A T T A T C	ASP GAT
S VAL GLY ASP VAL LEU ILE ARG GLU LYS LEU ASN A G T T G C C A T G T T T T A A T T C G A G A A A A T T A A T T C G A G A A A A T T A A T T C G A G A A A A T T A A A T T C G A G A A A A T T A A A T T C G A G A A A A T T A A A T T C G A G A A A A T T A A A T T C G A G A A A A T T A A T T C G A G A A A A T T A A T T C G A G A A A A T T A A T T C G A G A A A A T T A A T T C G A G A A A A T T A A T T C G A G A A A A T T A A T T C G A G A A A A T T A A T T C G A G A A A A T T A A T T C G A G A A A A T T A A T T C G A G A A A A T T A A T T A A T T C G A G A A A A T T A A T T C G A G A A A A T T A A T T C G A G A A A A T T A A T T A A T T C G A G A A A A A T T A A T T A A T T A A T T A A T T A A T T A A T T C G A G A A A A A T T T A A T T T T A T T T A T T T T A T	ARG C G C 820	N ASN ARG ALA GLU ILE LEU ILE GIN ILE ASN GLU ASP 1 TAATCGTGCTGAAATTTTAATTCAAATCAATGAAGATG 870 880 890	1 THR PHE LYS GLY ASN GLU SER VAL SER SER THR L AACTTTCAAGGGAACGAATCTGTTAGTAGCAGTACAT 930 940 950	LYS A A A 000	LEU GIN ALA ILE ARG ASP TYR TYR LEU ASN ASN GLY 1 CTGCAGGCAATTCGTGATTATTATTAAATAATGGCT 1050 1060 1060	ASN A A A T (1120
LEU	VAL GLY STAGGT	ILE A T T	ASN A A C	TRP TGG 1	ARG C G T (LEU CTA1
VAL G T T	VAL G T A	GLU GAA	G G G	TRP T G G	ILE ATTC	GLN C A G
C G A T C	A SER A A G T (810	3 ALA TGCT 870	: LYS CAAG 930	SER TCT1	ALA G C A A	VAL GTT 0.
G G C	ALA G C A 1	ARG C G T (870	語 TTC 99	ASP G A T 99	GN 7 CAGG 1050	ASP V G A T G 1110
VAL GTT	TYR T A T	ASN A A T	THR ACT	PRO C C T		THR ACG
PHE LYS TTAAA 740	GLU HIS 3 A G C A C 800	PRO ASN CAAAT 860	i 1	LEU GIN LTACAA 980	ASP G A T	THR LYS ACTAAA 1100
PHE T T T 740	GLU GAG 800	PRO C C A 860	SER T C A 920	LEU TTA 980	LYS A A A 1040	THR A C T 1100
G G G T	SER VAL LYS G T G T A A A A C	LEU CTG	ALA G C A	GLU GAA	GLU GAG	ILE A T C
ASN AAC	VAL G T A	THR ACG	LEU TTG	MET A T G	PHE TTC	GIN C A A
ASP ALA 3 A T G C T A 730	SER A G T 790	ASN A A T 850	ALA LYS LEU CAAAATTG 910	GLU GLN MET GLU 3AACAAATGGAA1 970	ALA GIN PHE GLU LYS ASP GCCAATTCGAGAAAGAT 1030 1040	LYS ALA GIN ILE A A G C A C A A A T C P 1090
	HE ALA LYS SER VAL LYS GLU HIS TYR ALA SER VAL GLY ARG TYR ASN ALA THR VAL GLU F T T G C C A A A A G A G C A C T A T G C A A G T G T C G C T A T A A C G C A A C G T T G A A C G T T G A A C G T T G A A C G T T G A A C G T T G A A C G T T G A C G T T G A C C A A C T G T G A C G T T G A C G T T G A C G T T G A C G T T G A C A C G T T G A C A C G T T G A C A C G T T G A C A C G T T G A C A C G T T G A C A C A A C A C A C A C A C A C A C	RO ILE VAL ASIN THR LEU PRO ASI CTATTGTCAATACGCTGCCAAA 850 850 860				YR ALA LYS ALA GIN ILE THR LYS THR ASP VAL GIN LEU ASN ASP GLU LYS THR LYS VAL 1 A T G C C A A A G C A C A A A T C A C T A A A C G G A T G T T C A G C T A A A T G A T G A A A A A A A G T T A 1090 1130 1140
LEU TTAC	ALA G C C	ILE	SP LYS ATAAA(EU GIN TACAA(LU GLY AAGGT(ALA G C C
A C	HE T T	RO L J	SP A T	EU T A	LU A A	YR A T

Ш
Z
\vdash
Z
0
\circ
$\widetilde{}$
Щ
<u>ത</u>
正

		38/8			
ILE G	HIS LEU ASN A	SP THR PHE ARG ARG SER ASP II.E ALA ASP VAL GLU ASN ALA II.E LYS ALA LYS I.EU GLY G	LYS T	ARG GIN LEU ARG P	ARG GIN GIN GIU G
TAG	SATTTAAATG	A T A C T T C C G C C G T A G T A T T G C A G A T G T A G A A A T G C A A T T A A A G C A A A A C T T G G G G	AAA	GCCAACTTCGCT	GACAACAAGAAG
1200	1260	1270 1270 1280 1290 1320	1380	1430 1440	1490 1500
SN VAL THR ILE ASP VAL ASN GLU GLY LEU GLN TYR ASP LEU ARG SER ALA ARG ILE ILE (A T G T A C T A A A T G A A G G T T T A C A G C T T C G T A G T G C A C G C A T T A T A G T C T T C G T A G T C A C G C A T T A T A G T C T T C G T A G T G T A T A T A G T T A T A G T C T T C G T A G T G T A T A T A T A G T C T T C G T A G T G T A T A T A G T G T A T A T	ALA GLU LEU GLU PRO LEU LEU SER ALA LEU HIS LEU ASN 1 GCCGAGCTTGAACCTTTACTTTCAGCATTACATTAAATG 1230 1240 1260	LEU CTTG	LU ARG GLY TYR GLY ASN THR THR VAL ASN SER VAL PRO ASP PHE ASP ALA ASN LYS A A C G A G G T T A C A C A C A G T A A T T C T G T A C T G A T T T G A C G A T G C A A A T T C T G T A C T G A B A T A A A A A A A A A A A A A A A A	LEU	HE GLU GLY ASN THR VAL SER ALA ASP SER THR LEU ARG GLN GLU MET ARG GLN GLU G TTGAAGGAAATACGTTGCTGATAGTACTTTACGTCAGGAAATGCGACAACAAGAAG 1450 1450 1460 1460 1470 1500
ARG C G C	HIS C A T	LYS A A A	ALA G C A	HR LEU ALA ILE THR PHE VAL VAL ASP ALA CLY ARG ARG LEU THR VAL ARG GIN LEU CATTAGCGATAACCTTGGTGGTGGTGGACGACGTTTAACTGTTCGCCAACTT 1390 1430	GIN C A A
ALA	LEU	ALA	ASP	ARG	ARG
G C A C	TTAC	G C A	G A T	C G C	C G A
1190	1250	1310	1370	1430	1490
ARG SER	SER ALA	LYS	ASP	VAL	MET
GTAGTO		A A A	G A C	G T T	A T G
ARG	SER	ILE	ME	THR	GU
CGT	TCA	A T T	TTT	ACT	GAA
LEU GIN TYR ASP LEU	LEU	ALA	THR VAL ASN SER VAL PRO ASP PHE ASP	ARG LEU THR VAL	SER THR LEU ARG GLN GLU METAGETACTTTACGTCAGGAAATGO
TTACAGTATGACCTT(CTTT	l G C A	ICAGTAATTCTGTACCTGATTTTGAC	CGTTTAACTGTTC	
1170 1180	1240	1300	1350	1420	
ASP G A C	PRO LEU : CTTTAC 12	ASN A A T	PRO C C T	ARG CGT	ARG CGT
TYR	PRO	GLU	VAL	ARG	LEU
T A T	C C T	GAA	G T A	C G A	TTA
GEN CAG	GLU GAA 30	VAL G T A	SER T C T	ALA GLY C T G G A C 1410	THR ACT
LEU C	LEU	ASP V	ASN S	ALA	SER :
TTAC	CTT	G A T G	AATT	G C T	A G T A
1170	123	1290	1350	141	1470
G G T	ALA GLU LEU GLU 3 C C G A G C T T G A 1 1230	ALA G C A G	val g t a	ASP GATO	ASP GAT
GLU	ALA	ILE	THR	VAL	ALA
GAAG	G C C	ATT (ACA	GTT(G C T C
ASN	SER	ASP	ASN THR	VAL	SER
A A T	T C T	G A T	IACACA	G T T	T C T
1160	1220	1280	1340	1400	1460
ASP VAL	GLY MET	ARG SER	ASN	PHE	ASN THR VAL
	GTATG1	GTAGTO	A A C	T T T	VATACCGTT1
ASP	GLY	ARG	G G T	THR	THR
G A T	GGT	C G T		ACC	ACC
ILE	ASN LEU GLY	ARG	ARG CLY TYR CLY	LEU ALA ILE THR PHE VAL	ASN
3 A T T C	AATCTGGGAC	C C C C	GAGGTTACGGTA	PTAGCGATAACCTTTGT1	A A A T
1150	1210	1270	1330	1390 1400	1450
VAL THR FIAACCA 11	LEU CTG	THR PHE CTTTCO	G G T	ALA G C G	GLY GGA
VAL G T A	LY ASN LEU GLY GLY MET SER GTAATCTGGGAGGTATGTCT(1210	THR ACT	ARG C G A	LEU TTA	HE GLU GLY TTGAAGGAA
A T. O	G T	SP A T A	LU A A	HR CAT	HE T T

GLU ASN T

呂

出

GLY GLY ASN ILE

CGGCAGTAAGTAGCTGGTACGAAAATGATTATGGTACGAGTGTCAATTTGGGTTATA

CCGAACCCTATTTACTAAAGATGGTGTAAGTCTTGGTGGAAATATTTTTTGAAAACT

GLY VAL SER LEU

ASP ASP

THR LYS

FIG.1E.(CONTINUED)

		39/	/82	
T _ 160	, c	C-4	A 7 G	₽
GLY F G G T T 1560	ASP G A T 16	G G T 16	GLY P GGGGG 1740	TYR
THR	VAL	TYR	THR	GLY
ACA	G T G	TAC	ACA	
ARG C G T	GLU GAA	G G T	GLY GGA	TEN.
ASP G A T 1550	ASP G A T 1610	11.E A T T 1670	TYR GIN THR SER ILE LYS GIN ASP ASN PHE LEU GLY THR GLY A NO FATCAAACAAGATAATTCTTGGGAACAGGGG17100 1720 1730 1730 1740	ASIN
LEU	ASN	GLY	PHE	VAL
T T A	A A T	GGT	T T C	
ARG	SER	PHE	ASN	SER
C G C	A G C	T T T	A A T	
S ILE	N GLY	ASN	N ASP	TH.
AATT	TGGT	A A C	A G A T	
1540	1600	660	1720	
LYS	ASN	ILE	GIN	GLY
A A A A	A A T	ATC	C A A	
G G A	ILE	SER A G T	LYS A A A	TYR
LEU TTA	PRO C C T	G G T	ILE A T T 0	ASP
GLU LI	ASP	THR	SER IL	ASIN
GAGT	G A T	A C G	AGTAT	
1530	159	165	1710	
VAL G T T	ILE	ASN A A C	THR ACA	LYS
LEU	ARG	ARG	GLN	開
T T A	C G A	C G T	C A A	
GLN	ASN	GLU	TYR	GLY
C A A	A A C	G A A	T A T	
1520	1580	1640	1700	
SER	GLU	LYS	SER	ALA
T C A	GAA	A A A	A G T	
ASN	VAL	VAL	ILE	ILE
A A T	G.T.T	GTC	A T C	
TYR	THR	LYS	GLY	SES.
3 T A T	A C A	A A A	F G G T	
1510	.570	.630	1690	
TRP	GUU	TYR	SER	VAL
TGG	GAA	TAT	AGT	
LY THR TRP TYR ASN SER GIN LEU VAL GLU LEU GLY LYS ILE ARG LEU ASP ARG THR GLY P GAACTTGGTATAATTCACAATTAGTTGAGTTAGGAAAATTCGCTTAGATCGTACAGGTT 1510 1520 1530 1550	HE PHE GLU THR VAL GLU ASN ARG ILE ASP PRO ILE ASN GLY SER ASN ASP GLU VAL ASP V TCTTCGAAACAGTTGAAAACCGAATTGATCCTATCAATGGTAGCAATGATGAAGTGGATG 1570 1570 1580 1590 1590	AL VAL TYR LYS VAL LYS GLU ARG ASN THR GLY SER ILE ASN PHE GLY ILE GLY TYR GLY T T C G T A T A T A A A G T C A A C G T A C A C G G G T A G T A T C A A C T T T G G T T A C G G T A 1 C G T A 1 6 G T	HR GLU SER GLY ILE SER TYR GLN THR SER ILE LYS GLN ASP ASN PHE LEU GLY THR GLY P CAGAGAGTGTATCAGTTATCAAGTATTAAACAAGATATTCTTGGGAACAGGGG 1730 1740	LA ALA VAL SER ILE ALA GLY THR LYS ASN ASP TYR GLY THR SER VAL ASN LEU GLY TYR T
LY G A	」 「 」 「	AL T C	HR C A	E

SUBSTITUTE SHEET

		4	10/82
SN V	HR T	зт.	R A
A T G	C C T	Г. G. A	A T A
1920	1980	2040	2100
T A I	TA (MA A 7	T. T. C. T. 2
SER	HIS	SE	ASN
A G	C A	T C	A A (
G G A	aty	GIN	TRP
	G G C	C A A	r g g
TYR T A T 1910	LEU F T A (ILE A T T (2030	GLY 3 G T 7 2090
THRACT	GLY	TYR	PHE
	3 G A	F A T i	[T T (
THR	VAL	LEU	SER
ACG	3 T A (FTA?	FCT7
THR SER SER ASN TYR LYS ARG THR THR TYR GLY SER ASN VACATCCTCTAACTATAAGCGTACGACTTATGGAAGTAATG	ASN GLU ASN ASN SER TYR TYR VAL GLY LEU GLY HIS THR T AATGAAAATAACTCCTATTATGTAGGATTAGGCCATACCT 1950 1960 1970 1980	ALA LEU CLU TYR ASN ARG ASN LEU TYR ILE GIN SER MET L GCTCTAGAATATAACCGTAATTTATATATTCAATGA 2010 2020 2030 2040	LYS THR ASN ASP PHE ASP PHE SER PHE GLY TRP ASN TYR A COAAATGACTTTGATTTTCTTTTGGTTGGAACTATA N 2070 2080 2080 2100
LYS	TYR	ARG	ASP
A A G	FAT		3 A T 7
TYR	SER	ASN	PHE
	T C C '	A A C (LTT(
ASN A A C	ASN A A C	TYR TATI	ASP 3 A C 1 0
SER 7	ASN 1	GU 1	ASN 7
TCTA	A A T A	GAAT	A A T G
1890	1950	2010	2070
SER	GLU	LEU	THR
T C C	GAAA	CTA	ACAI
~	ASN	ALA	LYS
	A A T	G C T (A A A I
ASP	VAL	PHE	
G A T	G T A	T T T	
1880	1940	2000	
SER A G T	PRO VAL CCTGT <i>1</i> 194(ASN A A C	G G C
LYS	班	SER	ASN
A A A	T T C	AGT	A A T
SER	GLY	11.E	GLY
CTCT	A G G T	A A T T	A G G T
1870	1930	1990	2050
ASN A A C	LEU TTA	LYS A A A	LYS A A A 2
YR ASP ASN SER LYS SER ASP	AL THR LEU GLY PHE PRO VAL	YR ASN LYS ILE SER ASN PHE	YS PHE LYS GLY ASN GLY ILE
A C G A T A A C T C T A A A A G T G A T	TTACTTAGGTTTCCCTGTA	ATAATAAATTAGTAACTTT	AATTTAAAGGTAATGGCATT
1870	1930 1940	1990 2000	2050 2060
YR	AL	YR	YS
A C	T T	A T	A A

ALA A TTACTATTCCAGGTTCTGATAACAAATACTACAAACTAAGTGCAGATGTACAGGGTTCT A C C C A T T A G A C A G A A T C A C C G C T G G G T T G T A T C T G C A A A A G C A T C T G C A G G A T A T G C A A GIN GLY NAL. ALA ASP 贸 ALA ALA LYS ALA SER B TRP VAL VAL SER TYR TYR LYS ASP ASN LYS ARG 段 ASP ARG ğ ASP PR 084 E PR 084 ¥

A C A G C C T T A A T A G A G G C T A T T T C C C A A C T A A A G G G G T T A A A G C A A G T C T T G G T G G A C G A G

G

ALA

VAL LYS

GLY.

THR LYS

8

黑

GLY

ARG

LEU ASN

SE

B

SUBSTITUTE SHEET

SAN GLY PHE GLY AND LINS ANG LED PRO PHE TYR GLY THE TYR THR ALA GLY GCTTTACCGTTCTATCAACTTATACAGCGGTGCGTTGGTTG	GATTAGAGAGCCAAGGICIIGAAAGACIIAACIIAITAIGGCAAAICAAGCCGIAIICGCG 2600 - 2600 - 2600 - 2600 - 2600 - 2610 - 2610 - 2620 - 2630 - 2630 - 264
	7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
THE AND THE ALL THE	⊈ → 5
GCA GCA AAC AAC CAA CAA ASP ATA ATA ARG	
ALA GLY GLY ILE CGGGTGGCATTG 2330 TYR ALA GLU HIS ATGCCGAACATG 2390 ALA ILE THR THR ALA ILE THR THR CAATCACAACC 2450 GLN ASN THR VAL CAAAACGGAAAAA 2550 SER ASP LYS AAATCAGATAAA 2570 SER SER ARG ILE	7630 2630
ALA GELN GELN GELN GELN GER	נ ב ר
FRO FHE TYR GIN THR TYR THR ALA GLY GLY ILE 2310 2320 2320 2320 SER ILE GLY FRO ASN ALA ILE TYR ALA GLU HIS SER ILE GLY FRO ASN ALA ILE TYR ALA GLU HIS 2370 2380 2380 2380 2390 SER ASP VAL ILE GLY GLY GLY ASN ALA ILE THR THR TCT GAT GT G GT G GT AAT G C AAT C A C A C T G 2430 PRO FHE VAL SER ASP LYS SER GLN ASN THR VAL C C A T T T G T G A T A A A A G C C A A A T A C A G T C C 2490 2500	द द
GIN THR TYR THR A A A C T T A T A C A G 2320 PRO ASN AIA ILE 2380 ILE GLY GLY ASN ITT G G T G G T A A T C 2440 SER ASP LYS SER A G T G A T A A A A G C 2500 ASN THR LYS TRP A T A C T A A A T G G I 2560 ASP TYR GLY LYS	ָ פר פר
THR ACT 2 ASN AAT GGT GGAT GAT ACT THR ACT THR ACT	ן א ד א ד
TYR GIN TYR GIN GLY PRO G G C C T I VAL ILE S T G A T T C TRP ASN	₹ 5
PHE TYR C 2310 2310 IIE GIY T G G G C 2370 ASP VAL ASP VAL PHE VAL TRP S T T G G P 2490 VAL TRP S T T G G P 2450 LEU PRO	ז כר ז נוס כי
PHE T T C T 2310 ILE G A T G 2430 PHE V T T G 2490 VAL 1 G T T T G 2550 LEU H	4 T T ,
PRO JER AGT A GER AGE AGE	ひなりゃ
TTA (GIY GGT) GGT AGT AGT AGT AGT AGT AGT AGT AGT AGT	3 A A C
LYS ARG LEU A G C G T T T A C 2300 ALA TYR GLY C T T A T G G T P 2360 LYS ILE SER A A G A T A A G T T 2420 VAL PRO THR 5 T A C C A A C T C 2480 ASP ALA ALA 3 A T G C G C A P 2540 VAL LEU LYS) I I 7
LYS T G C T T G T I T G A C T G A C T G A C T G A C	ָ פ פ
ASN A A A C A BHE T T T T G T A A T A T A T T C T A T T C T A T T C T A T T C	C A A
GTTTTGGAAACAAGCGTTTP GTTTTGGAAACAAGCGTTTP 2290 2300 2290 2300 2350 2350 2350 2350 2420 2420 2410 2410 2420 2420 2420 242	5 A 6
TTT TTTT A CG A GL A GL A GL A GL	A G A
SN GLY PHE GLY ASN LYS ARG LEU AT G G T T T G G A A C C A G C G T T T 2290 ER LEU ARG GLY PHE ALA TYR GLY C A T T A C G C G G T T T T G C T T A T G G 2350 SN GLY THR PHE ASN LYS ILE SER AT G G T A C T T T T A A T A A G A T A A G 2410 ER ALA GLU LEU ILE VAL PRO THE G T G C A G A A C T T A T T G T A C C A A C 2470 LY LEU PHE VAL ASP ALA ALA C C T C C T A T T T G T T G A T G C G C 2530 LY LEU GLU SER LYS VAL LEU LEU LYS LY LEU GLU SER LYS VAL LEU LYS LY LEU LEU LEU LEU LYS LYS LYS LYS LY LYS LY LEU LEU LEU LYS LYS LYS LYS LY LY LY LY LY LY LYS LY LYS LY LYS LY LY LYS LY LYS LY LYS LY LYS LY LYS LY LY LYS LY LYS LY LYS LY LY LYS LY L	T T W
	9

& 5 0	9 G
TYR T A T G 2700	GLY G G G G
SER T C T	ILE A T T
PHETTT	SER A G T
VAL G T A 2	PHE I T T ,
LEU P.T.G.	GIN C A A C
PRO	PHE l T C (
GIN TRP GIN SER PRO ILE GLY PRO LEU VAL PHE SER TYR A CAATGGCAATCTCCTATTGGACCATTGGTATTTCTTATG 2670 2680 2690	CELU ASN ASP ASP VAL GILU GIN PHE GIN PHE SER ILE GLY G TGAAAATGATGTCGAACAGTTCCAATTTAGTATTGGGG 10 2750 2760
ILE ATT (26	GLU 3 A A (27
PRO	VAL G T C (
SER LCT(ASP 3 A T (
GLN C A A T 2670	ASP 3 A T G 2730
TRP r g g (ASN A A T (
GIN C A A C	GLU 3 A A 1
., U O	TYR F A T (2720
GLY 3 G A 1	LYS
VAL GLY	LYS A A A
A G G T 0 2650	ILE A A T T 7 2710
LA SER THR GLY VAL GLY PHE CCTCTACAGGTGTCGGATT 2650 266	LA LYS PRO ILE LYS LYS TYR CTAAACCAATTAAAAATA 272
SER ICT1	LYS A A A (
LA C C J	LA LYS CTAAA

GCTCTTTCTAATAAATTGAACTTTTTCGTCATCAGAACTCAAAAACGACGTTCTGCC 2770 2780 2820 PHE ***

42/82

2850

SUBSTITUTE SHEET

cad15

(1-2949) (1-2953) (1-2984) (1-2989) (1-2974)	43/8	32	ACAGGACAGCTTTCCCTTTTAACCTTGAAATATTAGGGAAATTA	CTTTCC	TGA	aaaaggcattgaaaaaaaaggacagctttcccttttaaccttgaaaatattagggaaatta
cad15 minnad15 eagand15 pakd15 sb33d15	ᆏ	7	Н	[⊣	
1. cad15 3. minnad1 2. eagand1 4. pakd15 5. sb33d15	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

6 CTTaCTGGCGATTTGTCATTAAATAATTTAAGTGGGCCAATTTCTATTGCAAAAGGTGCTG	46 CTTCCTGGCGATTTGTCATTAATAATTTAAGTGGGCCAATTTCTATTGCAAAAGGTGCTG	62 CTTACTGGCGATTTGTCATTAAATAATTTAAGTGGGCCAATTTCTATTGCAAAAGGTGCTG		cttactggcgatttgtcattaaataatttaagtgggccaatttctattgcaaaaggtgctg	14/82	67 GCaCATCAGCAAATATTGGATTGGTGTATTTTTAAGTTTTATGGCACTGATTAGTGTAAA	107 GCcCATCAGCAAATATTGGATTGGTGTATTTTTAAGTTTTTATGGCACTGATTAGTGTAAA	123 GtGCATCAGCAAATATTGGATTGGTGTATTTTTTAAGTTTTATGGCATTGATTAGTGTAAA		gcgcatcagcaaatattggattggtgtattttttaagtttttatggca-tgattagtgtaaa
minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

gATTAC 	TTTAGGGATTATGAATTTATTTCCATTACCAGTATTAGATGGCGGTCATTTAGTTTTTTAA	TGAATTTATTTCCATTACCAGTATTAGATGCCGGTC	TTTAGGGATTATGAATTTATTT	TTATGAATTTATTTC	tttagggattatgaatttatttccATTACcagtattagatggcggtcatttagttttttta	45/82	ACAATGGAAGCTGTTAAAGGAAAACCTGTTTCTGAGCGGGTGCAAAGCATCTGTTATCGAA	ACAATGGAAGCTGTTAAAGGAAAACCTGTTTCTGAGCGGGTGCAAAGCATCTGTTATCGAA	ACAATGGAAGCTGTTAAAGGAAAA	ACAATGGAAGCTG1
	128	168	184	180		7	189	229	245	241
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15

acaatggaagctgttaaaggaaaacctgtttctgagcgggtgcaaagcatctgttatcgaa	gccAAGCTTAACGGTGTTTGCATTAATGATTTTTAATGATTTTAACGTCT	TTGGCGCAGCACTGTTATTAAGCTTAACGGTGTTTGCATTATTTAATGATTTTTACGTCT	TTGGCGCAGCACTGTTATTAAGCTTAACGGTGTTTGCATTATTTAATGATTTTTACGTCT	TTGGCGCAGCACTGTTATTAAGCTTAACGGTGTTTGCATTATTTAATGATTTTTACGTCT	TIGGCGCAGCACTGTTATTAAGCTTAACGGTGTTTGCATTATTTAATGATTTTTACGTCT Q ®	N ttggcgcagcactgttattAAGCTTAACGGTGTTTTGCATTATTTAATGATTTTTTACGTCT	ATAATTTATATAGGATACAATCGATGAAAAACTTCTAATCGCAAGTTTATTATTCGGTAC			
	7	250	290	306	302		52	311	351	367
consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15

sb33d15 consensus	363	ATAATTTATATAGGATACAATGGATGAAAAACTTCTAATCGCAAGTTTATTATTCGGTaC ATAATTTATATAGGATACAATGGATGAAAAACTTCTAATCGCAAGTTTATTATTCGGTaC
cad15	113	GACAACGACTGTGTTTGCCGCACCTTTTGTGGCAAAAGATATTCGTGTGGATGGTGTTCAA
minnad15	372	GACAACGACTGTGTTTGCCGCACCTTTTGTGGCAAAAGATATTCGTGTGGATGGTTCAA
eagand15	412	GACAACGACTGTGTTTGCCGCACCTTTTGTGGCAAAGATATTCGTGTGTGGGTGTTCAA
pakd15	428	GACAACGACTGTGTTTTGCCGCACCTTTTTGTGCCAAAAGATATTCGTGTGGTGTTCAAA
sb33d15	424	GACAACGACTGTGTTTGCCGCACCTTTTGTGGCAAAAGATATTCGTGTGGATGGTGTTCAA
consensus		GACAACGACTGTGTTTGCCGCACCTTTTGTGGCAAAAGATATTCGTGTGGATGTTCAA
cad15	174	174 GGTGACTTAGAACAACAAATCCGAGCAAGTTTACCTGTTCGTGCCGGTCAGCGTGTGACTG
minnad15	433	GGTGACTTAGAACAAAATCCGAGCAAGTTTACCTGTTCGTGCCGGTCAGCGTGTGACTC
eagand15	473	GGTGACTTAGAACAACAAATCCGAGCAAGTTTACCTGTTCGTGCCGGTCAGCGTGTGACTG

			4	18/8	2				5	
GGTGACTTAGAACAAATCCGAGCAAGTTTACCTGTTCGTGCtGGTCAGCGTGTGACTGTGACTGTGACTGTGACTGTGACTGTGACTGTGACTGTGACTGTGACTGTGACTGTGACTGTGACTGTGACTGTTTAGAACAAAATCCGAGCAAGTTTACCTGTTCGTGCGGGTCAGCGTGTGACTG	GGTGACTTAGAACAAAATCCGAGCAAGTTTACCTGTTCGTGCcGGTCAGCGTGTGACTG		Ø		Ø	Ø	ACAATGATGTGGCTAATATTGTCCGCTCTTTATTCGTAAGTGGTCGATTCGATGTGTGAA	AGCGCATCAAGAAGGCGATGTGCTTGTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	555 AGCGCATCAAGAAGGCGATGTGCTTGTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	
489		235	494	534	550	546		296	555	
pakd15 sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	

1C1/CA35/00

FIG.1F.(CONTINUED)

5 AGCGCATCAAGAAGGCGATGTGCTTGTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	611 AGCGCATCAAGAAGGCGATGTTGTTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	7 AGCGCATCAAGAAGGCGATGTGCTTGTTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	AGCGCATCAAGAAGGCGATGTGCTTGTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	GTTAAAAATCAAAGGTAACTCTGTTATTCCCACTGAAGCACTTAAACAAAACTTAGATGCTA	6 GTTAAAATCAAAGGTAACTCTGTTATTCCCACTGAAGCACTTAAACAAAACTTAGATGCTA @	GTTAAAATCAAAGGTAACTCTGTTATTCCCACTGAAGCACTTAAACAAAACTTAGATGCTA	2 GTTAAAATCAAAGGTAACTCTGTTATTCCCACTGAAGCACTTAAACAAAACTTAGATGCTA	8 GTTAAAATCAAAGGTAACTCTaTTATTCCacCTGAAGCACTaAAACAAAACTTAGATGCTA	GTTAAAATCAAAGGTAACTCTgTTATTCCcaCTGAAGCACTtAAACAAAACTTAGATGCTA	
595	61.	.09		357	616	929	672	899		
eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus	

ACGGGTTTAAAGTTGCCGATGTTTTAATTCGAGAAAATTAAATGAATTTGCCAAAAGTGT

418

cad15

ACGGGTTTAAAGTTGGCGATGTTTTAATTCGAGAAAAATTAAATGAATTTGCCAAAAGTGT	ACGGGTTTA	ACGGGTTTAAAGTTGGCGATGTTTTAATTCGAGAAAATTAAATGAATTTGCCAAAAGTGT		ACGGGTTTAAAGTTGGCGATGTTTTAATTCGAGAAAAATTAAATGAATTTGCCAAAAGTGT	Q 479 AAAAGAGCACTATGCAAGTGTAGGTCGCTATAACGCAACAGTTGAACCTATTGTCAATACGN	738 AAAAGAGCACTATGCAAGTGTAGGTCGCTATAACGCAACAGTTGAACCTATTGTCAATACG				AAAAGAGCACTATGCAAGTGTAGGTCGCTATAACGCAACaGTTGAACCTATTGAATAA
<i>L</i> 129	717	733	729		479	738	778	794	790	
minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

SUBSTITUTE SHEET

CATCATTAACTTTCAAGGGGAACGAATCTGTTAGTAGCAGTACATTACAAGAACAAATGGA

consensus

SUBSTITUTE SHEET

AAAGATLTGCAGLCAATTCGTGATTATTATTTAAATAATGGCTATGCCAAAGCACAAATLA

consensus

FIG. 1F. (CONTINUED)

ATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATAAAT	ATTACAACCTGATTCTTGGTGGAAATTATGGGGGAAATTAGAAGGTGCGCAATTCGAG	ATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATAAAT	ATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATTAAAATTTGAAGGTGCGCAATTCGAG		OPATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATAAAT	AAAGATTTGCAGTCAATTCGTGATTATTTAAATAATAGGCTATGCCAAAGCACAAATTA	AAAGATTTGCAGTCAATTCGTGATTATTATTAAATAATGGCTATGCCAAAGCACAAATTA	AAAGATTTGCAGTCAATTCGTGATTATTATTAAATAATGGCTATGCCAAAGCACAAATTA	AAAGATCTGCAGGCAATTCGTGATTATTATTTAAATAATGGCTATGCCAAAGCACAAATCA	AAAGATTTGCAGGCAATTCGTGATTATTTAATTAATAGCTATGCCAAAGCACAAATCA
662	921	961	977	973		723	982	1022		
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15 1038	sb33d15 1034

SUBSTITUTE SHEET

)2641A1_I_>

CTAAAACGGATGTTCAGCTAAATGATGAAAAAAAAAAGTTAATGTAACCATTGATGTAAA	CTAAAACGGATGTTCAGCTAAATGATGAAAAAACAAAAGTTAATGTAACCATTGATGTAAA	CTAAAACGGATGTTCAGCTAAATGATGAAAAACAAAAGTTAATGTAACCATTGATGTAAA	CTAAAACGGATGTTCAGCTAAATGAAAAAAAAAAAGTTAATGTAACCATTGATAAA	CTAAAGCGGATGTTCAGCTAAATGATGAAAAAAACAAAAGTTAATGTAACCATTGATGTAAA	CTAAAACGGATGTTCAGCTAAATGATGAAAAAAAAAAGTTAATGTAAACCATTGATGTAAAA	TGAAGGTTTACAGTATGACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTATGTCT	TGAAGGTTTACAGTATGACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTATGTCT	TGAAGGTTTACAGTATGACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTATGTCT	TGAAGGTTTACAGTATGACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTATGTCT	TGAAGGTTTACAGTATGACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTATGTCT	TGAAGGTTTACAGTATGACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTATGTCT
784	1043	1083		1095		845	1104	1144	1160	1156	
cad15	minnad15	eagand15	pakd15 1099	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

cad15	906	906 GCCGAGCTTGAACCTTTCAGCATTACATTTAAATGATACTTTCCGCCGTAGTGATA
minnad15	1165	GCCGAGCTTGAACCTTTACTTTCAGCATTACATTTAAATGATACTTTCCGCCGTAGTGATA
eagand15	1205	GCCGAGCTTGAACCTTTACTTTCAGCATTACATTTTAAATGATACTTTCCGCCGTAGTGATA
pakd15	1221	GCCGAGCTTGAACCTTTACTTTCAGCATTACATTTAAATGATACTTTCCGCCGTAGTGATA
sb33d15	1217	GCCGAGCTTGAACCTTTACTTTCAGCATTACATTTAAATGATACTTTCCGCCGTAGTGATA
consensus		GCCGAGCTTGAACCTTTACTTTACATTTAAATGATACTTTCCGCCGTAGTGATA & © © © © © © © © © © © © © © © © © ©
cad15	196	
minnad15	1226	TTGCAGATGTAGAAAATGCAATTAAAGCAAAACTTGGAGAACGCGGTTACGGTAGCGCAAC
eagand15	1266	TTGCAGATGTAGAAAATGCAATTAAAGCAAAACTTGGAGAACGCGGTTACGTAGCGAAC
pakd15	1282	
sb33d15	1278	TTGCAGATGTAGAAAATGCAATTAAAGCAAAACTTGGGGAACGAGGTTACGGTAACAAAC
consensus		TTGCAGATGTAGAAAATGCAATTAAAGCAAAACTTGGaGAACGcGGTTACGGTAACCAAC

GGTAAATTCAGTACCTGATTTTGATGATGCAAATAAACATTAGCGATAACCTTGTTGTT	nad15 1287 GGTAAATTCAGTACCTGATTTTGATGCAAATAAAACATTAGCGATAACCTTGTTGTT	GGTAAATTCAGTACCTGATTTTGATGATGCAAATAAAACATTAGCGATAACCTTGTTGTT	AGTAAATTCTGTACCTGATTTTGACGATGCAAATAAAACATTAGCGATAACCTTTGTTGTT	AGTAAATTCTGTACCTGATTTTGACGATGCAAATAAAACATTAGCGATAACCTTTGTTGTT	ggtaaattcagtacctgattttgatgatgcaaataaaacattagcgataaccttgttgtt (j	182	SACGACGTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTG	GATGCTGGACGATTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCGTTTTGTGCTG	GATGCTGGACGACGTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTG	GATGCTGGACGACGTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTG	GATGCTGGACGACGTTTAACTGTTCACCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTG	GATGCTGGACGACGTTTAACTGTTCgCCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTG
1028	1287	1327	1343	1339			1089	1348	1388	1404	1400	
cad15	minnad15	eagand15	pakd15	sb33d15	consensus		cad15	minnad15	eagand15	pakd15	sb33d15	consensus

ATAGCACTTTACGTCAGGAAATGCGCCAACAAGAAGGAACTTGGTATAATTCACAATTAGT	ATAGCACTTTACGTCAGGAAATGCGCCAACAAGAAGGAACTTGGTATAATTCACAATTAGT	ATAGCACTTTACGTCAGGAAA'TGCGCCAACAAGAAGGAACTTGGTATAATTCACAATTAGT	ATAGTACTTTACGTCAGGAAATGCGaCAACAAGAAGGAACTTGGTATAATTCACAATTAGT		ATAGCACTTTACGTCAGGAAATGCGCCAACAAGAAGGAACTTGGTATAATTCACAATTAGT U	ngagttaggaaaaattcgcttagatcgtacaggtttcttcgaaaacagtcgaaaatt	TGAGTTAGGAAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAAACAGTCGAAAACCGAATT	TGAGTTAGGAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAAACAGTCGAAATT	TGAGTTAGGAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAACAGTTGAAAACCGAATT	TGAGTTAGGAAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAACAGTTGAAAACCGAATT	TGAGTTAGGAAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAACAGTcGAAAACCGAATT
1150	1409	1449	1465	1461			1470	1510	1526	1522	
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15 1211	minnad15	eagand15	pakd15	sb33d15	consensus

1272 GATCCTATCAATGGTAGTAATGATGAAGTGGATGTCGTATATAAAGTCAAAGAACGTAACA	1531 GATCCTATCAATGGTAATGATGATGATGTCGTATATAAGTCAAAGAACGTAACA	1571 GATCCTATCAATGGTAGTAATGATGAAGTGGATGTCGTATAAAGTCAAAGAACGT	1587 GATCCTATCAATGGTAGCAATGATGAAGATGGATGTCGTATATAAAGTCAAAAGAACGTAACA	1583 GATCCTATCAATGGTAGCAATGATGAAGTGGATGTCGTATATAAAGTCAAAGAACGTAACA	GATCCTATCAATGGTAGtAATGAAGTGGATGTCGTATATAAAGTCAAAGAACGTAACA	1333 CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGAGTGGTATTAGTTATCAAG	1592 CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGAGTGTTTTATCAAG	1632 CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGTGGTATTAGTTATCAAG	1648 CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGAGTGGTATCAGTTATCAAACAAG	1644 CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGAGTGGTATtAGTTATCAAg	CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGAGTGGTATŁAGTTATCAAgCAAC
cad15	minnad15 1	eagand15 1	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensns

cad15 1394 TGTTAAACAAGATAATTTCTTGGGAACAGGGGCGGCAGTAAGTA	 TGTTAAA(TATTAAACAAGATAATTTCTTGGGAACAGGGGCGCCAGTAAGTA		TgTtaaacaagataatttcttgggaacaggggggggggtagtaagta	5 GATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTTACTAAAGATGGTGTA	1714 GATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTTACTAAAGATGGTGTAAA	GATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTTACTAAAGATGGTGTA	GATTATGGTACGAGTGTCAATTTGGGTTATACCGAACCCTATTTTACTAAAGATGTGTA		GATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTTACTAAAGATGGTGTAA
1394	1653	1693	1709	1705		1455		1754	1770	1766	
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15 1754	pakd15 1770	sb33d15	consensus

58/82

GTCTTGGTGGAAATGTTTTCTTTGAAAACTACGATAACTCTAAAAGTGATACATCCTCTAA	GTCTTGGTGGAATGTTTTCTTTGAAACTACGATAACTCTAAAAAGTGATACATCCTCTAA	GTCTTGGTGGAAATGTTTTTTTTTTTTTTTTTTTTTTTT	GTCTTGGTGGAAATATTTTCTTTGAAACTACGATAACTCTAAAAGTGATACATCCTCTAA	GTCTTGGTGGAAATGTTTTCTTTGAAAACTACGATAACTCTAAAAGTGATACATCCTCTAA	GTCTTGGTGGAAATGTTTTCTTTGAAACTACGATAACTCTTAAAAGTGATACATCCTCTAA U		CTATAAGCGTACGACTTACGGAAGTAATGTTACTTTAGGTTTCCCTGTAAATGAAATAAC	CTATAAGCGTACGACTTACGGAAGTAATGTTACTTTAGGTTTCCCTGTAAATGAAAATAAC	CTATAAGCGTACGACTTACGGAAGTAATGTTACTTTAGGTTTCCCTGTAAATGAAAATAAC		CTATAAGCGTACGACTTATGGAAGTAATGTTACTTTAGGTTTCCCTGTAAATGAAAATAAC	CTATAAGCGTACGACTTAcGGAAGTAATGTTACTTTAGGTTTCCCTGTAAATGAAAATAAC
1516	1775	1815	1831	1827		 	1577	1836	1876	1892	1888	
cad15	minnad15 1775	eagand15	pakd15	sb33d15	consensus	1	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

TCCTATTATGTAGGATTAGGTCATACCTATAAAATTAGTAACTTTGCTCTAGAATATA	TCCTATTATGTAGGATTAGGTCATACCTATAATAAAATTAGTAACTTTGCTCTAGAATATA	TCCTATTATGTAGGATTAGGTCATACCTATAATAAAATTAGTAACTTTGCTCTAGAATATA	TCCTATTATGTAGGATTAGGCCATACCTATAATAAAATTAGTAACTTTGCTCTAGAATATA		TCCTATTATGTAGGATTAGGtCATACCTATAATAAAATTAGTAACTTTGCTCTAGAATATA O	O/8	ACCGTAATTTATATTCAATCAATGAAATTTAAAGGTAATGGCATTAAAACAAATGACTT N	ACCGTAATTTATATTCAATCAATGAAATTTAAAGGTAATGGCATTAAAACAAATGACTT	ACCGTAATTTATATTCAATCAATGAATTTAAAGGTAATGGCATTAAAAACAAATGACTT	ACCGTAATTTATATTCAATCAATGAATTTAAAGGTAATGGCATTAAAACAAATGACTT		ACCGTAATTTATATATTCAATCAATGAAATTTAAAGGTAATGGCATTAAAACAAATGACTT
	1897	1937	1953	1949			1699	1958	1998	2014	2010	
cad15 1638	minnad15	eagand15	pakd15	sb33d15	consensus		cad15	minnad15	eagand15	pakd15	sb33d15	consensus

TGATTTTTCCTTTGGTTGGAACTATAACAGCCTTAATAGAGGCTATTTCCCAACTAAAGGG	TGATTTTTCTTTTGGTTGGAACTATAACAGCCTTAATAGAGGCTATTTCCCAACTAAAGGG	TGATTTTTCTTTTGGTTGGAACTATAACAGCCTTAATAGAGGCTATTTCCCAACTAAAGGG	TGATTTTTTTT	TGATTTTTTT	TGATTTTTCTTTTGGAACTATAACAGCCTTAATAGAGGCTATTTCCCCAACTAAGGG O			_			GTTAAAGCAAGTCTTGGTGGACGAGTTAC
1760	2019	2059	2075	2071		1821	2080	2120	2136	2132	
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

cad15 1882	1882	TAAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAGATCACCTCTGGGTTGTATCTGC
minnad15 2141		TAAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAGATCACCTCTGGGTTGTATCTGC
eagand15	2181	TAAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAGATCACCTCTGGGTTGTATCTGC
pakd15 2197	2197	TAAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAGATCACCGCTGGGTTGTATTTGCCTGC
sb33d15	2193	TAAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAGATCACCtCTGGGTTGTATCTGC
consensus		TAAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAGATCACCtCTGGGTTGTATCTGC
cad15	1943	AAAAGCATCTGCAGGATATGCAAATGGTTTTGGAAACAAGCGTTTACCGTTCTATCAAACT
minnad15 2202		AAAAGCATCTGCAGGATATGCAAATGGTTTTGGAAACAAGCGTTTACCGTTCTATCAAACT
eagand15	2242	AAAAGCATCTGCAGGATATGCAAATGGTTTTGGAAACAAGCGTTTACCGTTCTATCAAACT
pakd15 2258	2258	
sb33d15	2254	AAAAGCATCTGCAGGATATGCAAATGGTTTTGGAAACAAGCGTTTACCGTTCTATCAAACT

AAAAGCATCTGCAGGATATGCAAATGGTTTTTGGAAACAAGCGTTTACCGTTCTATCAAACT	TATACAGCGGGTGGCATCGGTTCATTACGTGGTTTTGCTTATGGTAGTATTGGACCTAACG	TATACAGCGGGTGGCATCGGTTCATTACGTGGTTTTGCTTATGGTAGTATTGGACCTAACG	TATACAGCGGGTGGCATCGGTTCATTACGTGGTTTTGCTTATGGTAGTATTGGACCTAACG	TATACAGCGGGTGGCATTGGTTCATTACGCGGTTTTTGCTTATGGTATTGGCCTAATG	TATACAGCGGGTGGCATTGGTTCATTACGCGGTTTTTGCTTATGGTAGCATTGGGCCTAACG	TATACAGCGGGTGGCATcGGTTCATTACGtGGTTTTGCTTATGGTAGtATTGGACCTAAcG N	CAATTTATGCCGAATATGGTAATGGTAGTGGTACTGGTACTTTTAAGAAGATAAGTTCTGA	CAATTTATGCCGAATATGGTAGTGGTACTGGTACTTTTAAGAAGATAAGTTCTGA	CAATTTATGCCGAATATGGTAGTGGTACTGGTACTTTTAAGAAGATAAGTTCTGA		
	2004	2263	2303	2319	2315		2065	2324	2364	2380	
consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	

sb33d15 2376 CAATTTATcaaGgtCAaaaTAAT	CAATTTATgccGaatAtggTAATggtagtggtactggtactTTTAAgAAGATAAGTTCTGA	TGTGATTGGTGGTAATGCAATCGCTACAGCTAGCGCAGAGTTAATTGTGCCAACTCCATTT		TGTGATTGGTGGTAATGCAATCGCTACAGCTAGCGCAGAGTTAATTGTGCCAACTCCATTT	TGTGATTGGTGATGCAATCaCaACtGCGAGtGCAGAacTtATTGTaCCAACTCCATTT O		TGTGATTGGTGGTAATGCAATCgCtACaGCtAGcGCAGAgtTaATTGTgCCAACTCCATTT	
2376		2126	2385	2425	2429	2422		
sb33d15	consensus	cad15 2126	minnad15	eagand15 2425	pakd15 2429	sb33d15 2422	consensus	

GTGAGCGATAAGAGCCAAAATACGGTCCGAACCTCCTTATTTGTTGATGCGGCAAGTGTTT GTGAGCGATAAGAGCCAAAATACGGTCCGAACCTCCTTATTTGTTGATGCGGCAAGTGTTT GTGAGCGATAAGAGCCAAAATACGGTCCGAACCTCCTTATTTGTTGATGCGGCAAGTGTTTT 2187 2446 2486 cad15 eagand15 minnad15

GTGAGTGATAAAAAAAAATACAGTCCGAACCTCCCTATTTGTTGATGCGGCAAGTGTTT		GTGAGCGATAAgAGCCAAAATACgGTCCGAACCTCCtTATTTGTTGATGCGGCAAGTGTTT	GGAATACTAAATGGAAATCAGATAAAAATGGATTAAGAGCGATGTATTAAAAAGATTGCC	GGAAATCAGATAAAAATGGATTAGAGAGCGATGTATTAAAAAGATTGCC	GGAATACTAAATGGAAATCAGATAAAAATGGATTAGAGAGCGATGTATTAAAAAGATTGCC N	GGAATACTAAATGGAAATCAGATAAAAATGGATTAGAGAGCAAGGTCTTGAAAGACTTACC		GGAATACTAAATGGAAATCAGATAAAAATGGATTAGAGAGCgAtGTaTTaAAAAagaTTgCC	TGATTATGGCAAATCAAGCCGTATTCGCGCCTTCTTACAGGTGTTCGCAATTCAAATCAAATTATATTCAAAATCAAAATCAAAATTAATTCGCAAAATCAAAATTAATT		
2490	2483		2248	2507	2547	2551	2544		2309	2568	
pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	

2608 TGATTATGGCAAATCAAGCCGTATTCGCGCCTCTACAGGTGTCGGATTCCAATGGCAATCT	2612 TGATTATGGCAAATCAAGCCGTATTCGCGCCTCTACAGGTGTCGGATTCCAATGGCAATCT	2605 cGATTATGGCAAATCAAGCCGTACTCGCGCCTCTACAGGTGTCGGATTCCAATGGCAATCT	tGATTATGGCAAATCAAGCCGTAtTCGCGCCTCTACAGGTGTCGGATTCCAATGGCAATCT	2370 CCTATTGGCCATTGGTATTCTCTTATGCCAAACCAATTAAAAAATATGAAAATGATGATGO	2629 CCTATTGGGCCATTGGTATTCTCTTTATGCCAAACCAATTAAAAATATGAAAATGATGATGATGATGATGATGATG	2669 CCTATTGGGCCATTGGTATTCTCTTATGCCAAACCAATTAAAAAATATGAAAATGATGATGATGATGAATGA	2673 CCTATTGGACCATTGGTATTTTCTTATGCTAAACCAATTAAAAAATATGAAAATGATGATGATG	2666 CCTAGTGGACCAGTGGTATTTTCTTATGCTAAACCAATTAAAAAATATGAAAATGATGATGATG	CCTAtTGGgCCAtTGGTATTcTCTTATGCcAAACCAATTAAAAAATATGAAAATGATGATG	
eagand15	pakd15 2612	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus	

minnad15 2690 TCGAACAGTTCCAATTTAGTATTGGAGGTTCTTTCTAATAAATTGAACTTTTTTTT

67/82

FIG.1F.(CONTINUED)

	-				o ATTTAATTAAGGATATTTATCAAATGAAAAACATCGCAAAAGTAACCGCACTTGCTTTAGG © O	4 TATTGCACTTGCTTCAGGCTATGCTTCCGCTGAAGAAAATTGCTTTCATTAATGCAGGT					TaTTGCACTTGCTTCAGGCTATGCTtCcGCTGAAGAAAAATTGCTTTTCATTAATTGCTTTTTCATTAATTGCTTATGCTATGCTATGCTATGCTATGCTATGCTATGCT
2553	2812	2852	2856	2849		2614	2873	2913	2917	2910	
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

101/0:150/0000

FIG.1F.(CONTINUED)

cad15 2675 atattTTTcaAcatCacccagatcgccaagcggtagcagataaacttgatgctgaatttaa	minnad15 2934 TATAnTTTnCAAggCnaagg	 eagand15 2973 TATALTTTCAA 	 pakd15 2977 TTATATTTTCAa 	 Sb33d15 2970 TTATA	is ttat-ttttcaaa-cgatcgccaagcggtagcagataaacttgatgctgaatttaa
cad	minnad1	eagand1	pakd1	sb33d1	consensus

acctgtagctgagaaattagcagcaagcaaaaaaaagaagttgatgataaaattgctgctgct cad15 2736

69/82

minnad15 2954

eagand15 2985

pakd15 2990

sb33d15 2975

consensus

acctgtagctgagaaattagcagcaagcaaaaaaagaagttgatgataaaattgctgctgct

FIG.1F.(CONTINUED)

cad15 2797 cgtaaaaaagtagaagcaaaagttgcggctttagaaaaaagatgcacctcgcttacgtcaag

minnad15 2954

2985 eagand15

sb33d15 2975 pakd15

2990

consensus

 $\bigcup_{\text{cgtaaaaaagtagaaggcaaaagttgcggctttagaaaaaaagatgcacctcgcttacgtcaag}$ ctgatattcaaaaacgccaacaggagattaataaattaggtgcggctgaagatgctgaatt 2858 cad15

minnad15 2954

2985 eagand15

pakd15 2990

2975 sb33d15

consensus

ctgatattcaaaaacgccaacaggagattaataaattaggtgcggctgaagatgctgaatt

SUBSTITUTE SHEET

71/82

FIG.1F.(CONTINUED)

cad15 2919 acaaaaattaatgcaagaacaagataaaaa

minnad15 2954 eagand15 2985

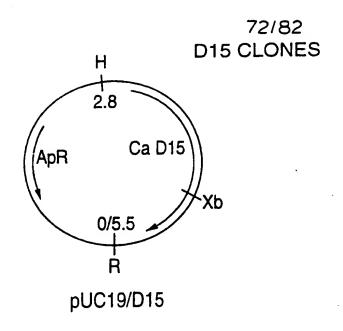
pakd15 299

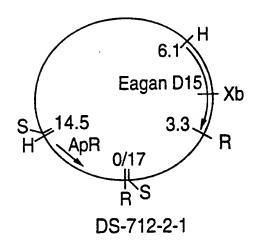
sb33d15 2975

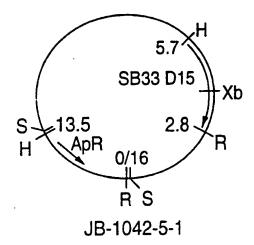
consensus

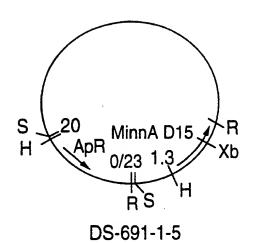
acaaaaattaatgcaagaacaagataaaaaa

WO 94/12641 PCT/CA93/00501









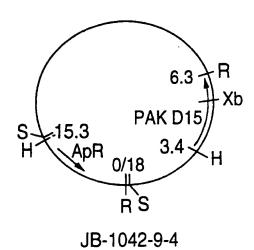


FIG.2.

SUBSTITUTE SHEET

D15 SEQUENCE COMPARISON

MKKLILIASILIEGITTIVEAAPEVAKDIRVDGVQEDLEQQIRASLEVRAGQRVIINIDVANIVRSLEVSGREDDVKAHQEGDVLAVSVVAKSILSDVKIKGN	පු
	Eagan
	MinnA
	SB33
	PAK
SVI PTEALKONI DANGFKVGDVLIREKI NEFAKSVKEHYASVGRYNATVEPIVNIT PANRAETLIQINEDDKAKLASLITFKGNESVSSSTLQEQMELQPD	පු
	Eagan
	MirmA
I	SB33
	PAK
SWWKIWCHECZĄFEKDIĄSIRDYYIANGYAKAQITKIDVQIADEKIKVANTIDANECIĄYDIRSARIIGALGCASAELEPLISALHIADIFPRRSDIAD	පු
	Eagan
	MinnA
	SB33
	PAK
VENATKAKI GERGYGSATVNSVPDFDDANKTI ALTILAVDAGRRUTVRQI RFEGANTVSADSTI RQEMRQQEGTWYNSQLVELGKI RLDKIGFFETVENRID	පු
	Eagan
	SB33
F.	PAK

73/82

SUBSTITUTE SHEET

									7	4/82	2					
යි	Eagan	MirmA	SB33	PAK	පු	Eagan	MirmA	SB33	PAK	පී	Eagan	MirmA	SB33	PAK	Ca Eagan MirmA SB33	PAK
PINGSNDENDAVYKVKERNIGSINFGIGYGIESGISYQASVKQINFLGIGAAVSIAGIKNDYGISANLGYTEPYFTKIGYSIGANVFFENYINSKSDISS				T.TI.	NYKRITYGSNVILGEPANBANSYYVGLGHIYNKISNFALEYNRALYIQSMKFKGAGIKINDFDFSFGMAYNSIARGYFPIKGVKASLGGRVTIFGSIDAKY					YKLSADVQGFYPIJRDHLWVVSAKASAGYANGFGNKRIJPFYQTYTPAGGIGSIJRGFAYGSIGBNATYAEYGNGSGTGTFKKISSDVIGGNALATASAELIV				R	PIPFVSDKSQNIVRISLFVDAASVANIRVIKSDRNGLESDVLKRLPDYGKSSRURASIGVGFQAQSPIGPLVFSYAKPIKKYENDJOVEQFQFSIGGSF** ** ** N. D.	**

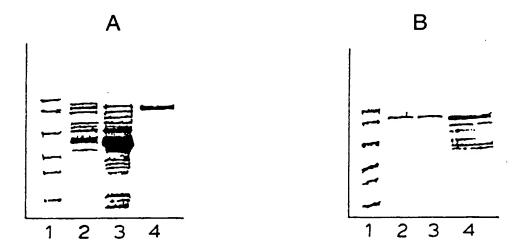
10 34/12041

75/82 Construction of plasmid expressing SB33 D15 5.7 D15 **ApR** JB-1042-5-1 |3.5 2.8 |ApR 0/16 pUC 2.8 ↓R/H RH H BsrF I **ApR** D15 \ Xb pUC BsrF I pRY-60-1 BsrF I Nde RH BsrF I/R R/Nde Ŕ Nde-BsrF I oligos Nde ∠BsrF I ApR D15 ApR pUC BsrF I DS-860-1-1 pT7-7 Bg Bg/Nde R/Rg Nde Nde/R Bg H/ Nde **D15** 'ApR DS-880-1-2 Xb

FIG.4.

WO 94/12641 PCT/CA93/00501

76/82 PURIFICATION OF D15 FROM A NON-TYPEABLE HAEMOPHILUS INFLUENZAE STRAIN 30



PROTEIN STAIN

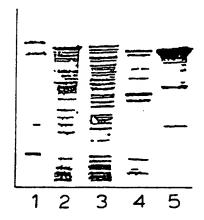
WESTERN BLOT

- 1. Low MW markers
- 2. Strain 30
- 3. Native D15 crude extract
- 4. D15 after anti-D15 affinity chromatography

FIG.5.

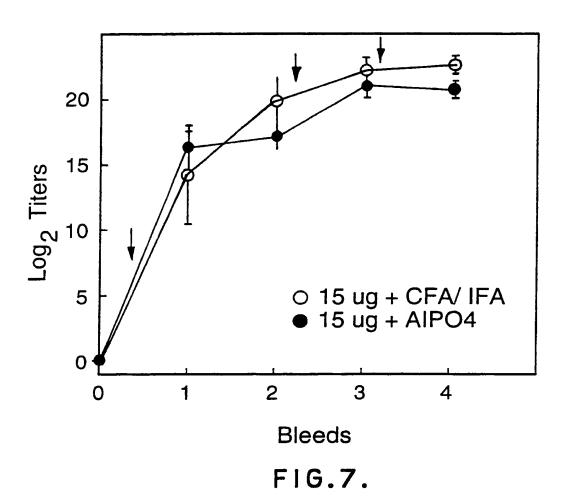
77/82

PURIFICATION OF FULL LENGTH RECOMBINANT D15



- 1. Protein M.W. Markers
- 2. Lysate of E. coli expressed rD15
- 3. Soluble protein in Tris-HC1 buffer extract
- 4. Soluble proteins in Tris/Triton X-100/ EDTA extraction buffer
- 5. rD15 inclusion bodies

FIG.6.



O 94/12641 PC1/CA93/005



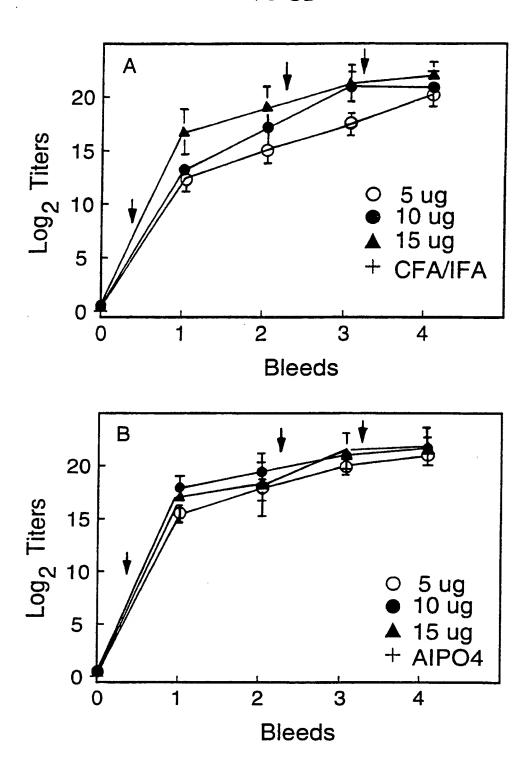
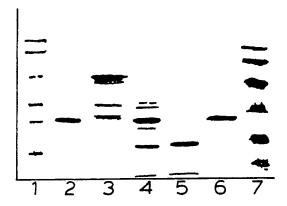


FIG.8.

80/82

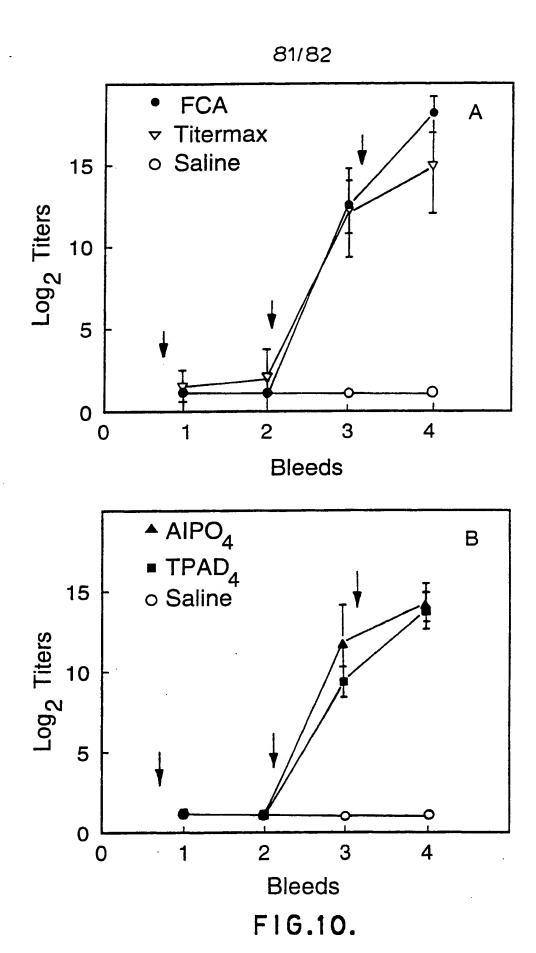
PURIFICATION OF TRUNCATED D15 FROM D15-GST FUSION PROTEIN



- 1. Prestain low MW markers
- 2. GST standard
- 3. GST-(D15 fragment) fusion protein
- 4. Fusion protein cleaved by thrombin
- 5. rD15 fragment
- 6. GST
- 7. Low MW markers

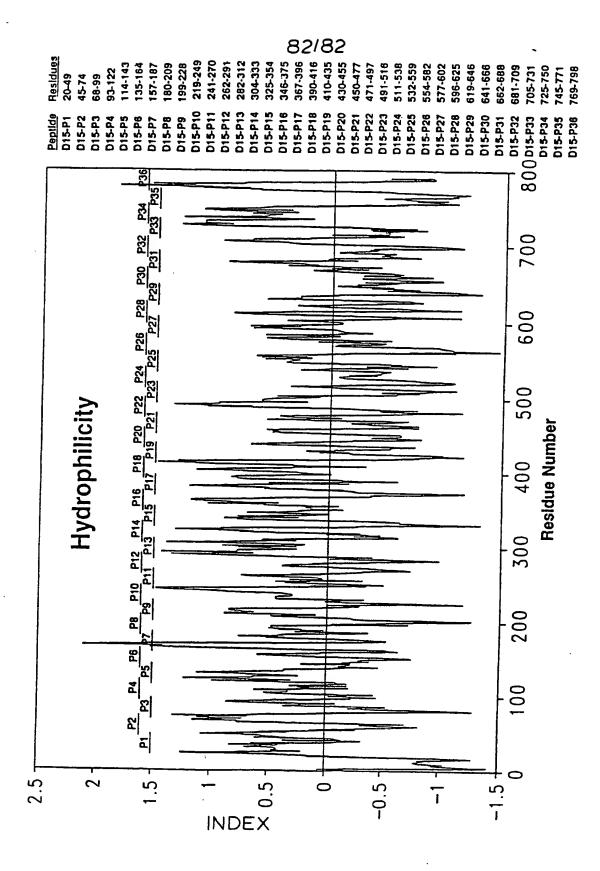
FIG.9.

PCT/CA93/00501



SUBSTITUTE SHEET





HATEVILL TIONAL SEAKCH KELOK!

-2-

tntr mal Application No
PCT/CA 93/00501

stegory '	ction) DOCUMENTS CONSIDERED TO BE RELEVANT Citatum of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
J/		ACTIONAL TO CLAIM: No.
,	(CONNAUGHT LABORATORIES LIMITED et al.) 16 May 1991 (16.05.91), claims.	12,16, 20,22
A	EP, A1, 0 281 673 (THE RESEARCH FOUNDATION OF STATE UNIVERSITY OF NEW YORK) 14 September 1988 (14.09.88), claims.	1,4,6, 12,18, 22,27
	US, A, 5 013 664 (BRODEUR et al.) 07 May 1991 (07.05.91), abstract.	1,27
	·	
	·	

zum internationalen Recherchenbericht über die internationale

Patentanneldung Nr.

to the International Search Report to the International Patent Application No.

au rapport de recherche inter-national relatif à la demande de brevet international n'

PCT/CA 93/00501 SAE 82214

angeführten Patentdokumente angegeben. Diese Angaben dienen nur zur Unterrichtung und erfolgen shoe Gewähr.

In diesem Anhang sind die Mitglieder der Patentfamilien der im obenge- members relating to the patent documents members de la famille de brevets relatifs aux documents de brevets relatifs aux documents de brevets national search report. The Office is in no way liable for these particulars which are given merely for the purpose of information.

relatifs aux documents de brevets cités dans le rapport de recherche inter-national visée ci-dessus. Les reseignements fournis sont donnés à titre indicatif et n'engagent pas la responsibilité de l'Office.

Is Recherchenbericht angeführtes Patentdokusest Patent document cited in search report Document de brevet cité dans le rapport de recherche	Datum der Veröffentlichung Publication date Date de publication	Patentfamilie Ver Patent family member(s) Membre(s) de la	Datum der Veröffentlichung Publication date Date de publication		
EP A2 378929	25-07-90	CA AA 2006587 23 CA AA 2006587 23 EP A3 378929 03 JP A2 2283286 20	2-03-89 2-06-90 3-06-90 1-08-90 0-11-90 2-02-89		
WO A1 9106652	16-05-91	EP A1 500576 02 GB A0 8924473 20	1-05-91 2-09-92 0-12-89 1-03-93		
EP A1 281673	14-09-88	DK AO 6020/87 17 DK A 6020/87 19	2-12-92 7-11-87 7-05-88 0-06-89		
US A 5013664	07-05-91	keine - none - rien			